Roger Penrose
Collected Works

VOLUME 4

1981–1989

OXFORD SCIENCE PUBLICATIONS

OXFORD
UNIVERSITY PRESS

www.oup.com

ISBN 978 0 19 921939 1

ROGER PENROSE
COLLECTED WORKS
VOLUME 4
1981–1989

常州大学山书馆
藏书章

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Roger Penrose 2011
The moral rights of the authors have been asserted
Crown copyright material is reproduced under Class Licence
Number C01P0000148 with the permission of OPSI
and the Queen's Printer for Scotland
Database right Oxford University Press (maker)
First published 2011
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Control Number: 2010926298
Typeset by Newgen Imaging Systems Pvt Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire
ISBN 978-0-19-921939-1 (Vol. 4)
978-0-19-921944-5 (Set)
1 3 5 7 9 10 8 6 4 2

PREFACE

The topics addressed, in these *Collected Works*, vary over quite an extensive range, the boundaries between their areas being sometimes quite blurred. It would have been hard to classify all these articles and arrange them according to individual topics. Moreover, my own mode of working has been, on the whole, to keep most of my different scientific or mathematical interests running concurrently. Accordingly, it has seemed appropriate not to attempt to group the papers together into separate tidy categories. To do so would have provided a somewhat misleading picture of my overall mode of thinking. Instead, the ordering of articles that has been adopted is essentially chronological, at least with regard to the date of publication. The resulting somewhat higgledy-piggledy appearance that we find in the sometimes abrupt juxtaposition of different topics is, perhaps, fairly illustrative of my actual mode of working.

There are two main kinds of drive that underlie these works. On the one hand, some of the articles have arisen mainly from purely mathematical considerations, which can be ultimately aesthetic or just matters of simple curiosity. On the other hand, motivations for many others of the articles came primarily from a strong desire to discover more about the underlying principles governing the workings of the physical world. Yet, one of the deep mysteries about what has been discovered about these principles is the great breadth of intrinsically appealing mathematics that is found to underlie the actual operation of the physics of the universe, so these two drives are found to relate closely to one another. The reader may perhaps note the distinct geometrical flavour of much of my own mathematics, evidencing the strong appeal that geometry exerts on me, an appeal that is not infrequently intertwined with desire for algebraic elegance. Indeed, I have found it remarkable how much of the workings of our universe appear to depend upon sophisticated and powerful geometry and algebra.

My own very geometrical perspectives, as opposed to purely computational ones, are perhaps made more obvious in the later volumes of this series than in the early ones, but this appearance is deceptive. Very early in my career as a Cambridge graduate student, I had developed a graphical notation which enabled me to carry out many kinds of calculation in an entirely diagrammatic way. For a long time I was rather secretive of this procedure, and in my earlier writings would take care to remove all trace of these visual origins. I do not doubt that my predilections for such geometric or algebraically elegant forms of understanding have influenced the specific choices of direction in which my endeavours towards furthering our understandings of the physics of the world have taken.

The focus of these directions has become more clarified as time has moved on, and has narrowed down to such a degree that there has emerged a somewhat disturbing repetitiveness in many of the later articles. A good number of these provide rather non-technical accounts. An example of such repetitiveness occurs with my descriptions of gravitational collapse to a black hole. There is, however, a gradually developing outlook on this issue, leading to an emphasis on the fundamental contrast in character between the space-time singularities that inevitably appear within black holes and the space-time singularity that constitutes the big bang, this contrast being central to the existence of a second law of thermodynamics in our universe. The same kind of repetitiveness applies to my frequently repeated claim of a need for a fundamental change in the basic laws of quantum theory, where only towards the end do my suggestions become particularly specific. Likewise, many mildly differing accounts are provided outlining the formalism and the basic ideas that underlie twistor theory. The later ones lead up to various viewpoints on the 'googly' problem, which has constituted the main obstruction to the finding of a full understanding of general relativity in twistor terms. Despite such repetition, it is my hope that there is a sufficient variety of emphasis and of historical development in such articles for it to have been worth while to include them all here.

Several colleagues have collaborated with me in writing the works presented here, and I can mention but only a few. Foremost, I have collaborated long and particularly fruitfully with Ezra T. Newman, especially with regard to the asymptotics of general relativity, and also with Wolfgang Rindler in the development of the formalism of 2-spinor calculus. Other collaborators who have played important roles in helping to develop the mathematics of discrete physics and twistor theory have been Erwin Kronheimer, Michael Eastwood, Ronny Wells, and Denny Hill. Among my collaborators, providing very significant inputs, who were former students in my research group, are Andrew Hodges, Paul Tod, Lane Hughston, George Sparling, Richard Ward, Nick Woodhouse, and Lionel Mason, but there were many others. Although not playing roles as actual collaborators, fundamental influences in my research have come from Dennis Sciama and Michael Atiyah.

Roger Penrose
April 2010

BIOGRAPHY

Early years and influences

I was born in Colchester, Essex, UK on 8 August 1931. My father, Lionel Sharples Penrose, FRS, was a highly talented and intelligent man, with many interests. He, himself, had come from a Quaker family, his own father (James Doyle Penrose) having been a professional artist and his mother (Josephine Peckover) was from a well-to-do Quaker banking family. Lionel's primary professional interest was in the causes of hereditary mental disease. He was Galton Professor of Human Genetics at University College London from 1945 and, following retirement, Director of the Kennedy–Galton Institue, St Albans. He had a very considerable artistic talent and a keen interest in music, especially Mozart, but also Bach and other classical composers, dabbling a little in composition himself, particularly when there was a puzzle aspect in what was involved (where he once wrote a piece that could be played equally well upside down as right-way up). He liked making puzzles of various kinds in wood, and spent much of his later life designing wooden models illustrating self-reproduction. He was recognized as a fine composer of chess problems. He had a definite fondness for mathematics and was an innovative statistician. He loved explaining matters of science and mathematics. On one occasion, when I told him that my mathematics teacher at school had informed us that he would be explaining calculus the next day, Lionel immediately took me aside so as to be sure to be the first to teach me about the beauties of calculus. This made a great impression on me. Yet he was an extremely difficult person to communicate with on personal matters.

My mother, born Margaret Leathes, was a charming and caring woman with a very subtle mind. She had been Head Girl at Bedales school, and became medically qualified, but found she was not able to practice after her marriage. After Lionel died, in 1972, she married Max Newman FRS, a fascinating man, who had been Professor of Mathematics in Manchester and, as we later learned, a key figure at Bletchley Park during the war. Margaret's father, John Beresford Leathes, FRS, had been a distinguished professor of physiology, and her mother, a concert pianist, born Sara Mara Natanson from a Jewish family in Latvia, had disowned her family (then in St Petersberg), becoming completely secretive about her origins, and even about her maiden name. Margaret enjoyed geometry at school, but did not

communicate much on this topic with me. My older brother Oliver was very precocious, and I learnt a lot of physics and mathematics from him as I was growing up. He later became a distinguished mathematical physicist (statistical mechanics) and FRS. My younger brother Jonathan excelled at games, particularly chess, and later became British Chess Champion a record 10 times, and then a top correspondence chess player. My much younger sister Shirley Victoria, now Shirley Hodgson, became a highly regarded Professor of Cancer Genetics.

This atmosphere of science, mathematics, art, and music, and of puzzles and games, had been an important part of my upbringing. During the war years, the family lived in London, Ontario, Canada, where Lionel became Director of Psychiatric Research at the Ontario Hospital. We returned to England in 1945, where I attended University College School, London. As an undergraduate, I studied Mathematics at University College London, from 1951 to 1954; then I went to St John's College Cambridge to do research in algebraic geometry, initially under William Hodge and subsequently John Todd. In my same year, under Hodge, was Micheal Atiyah, who was an inspirational, though somewhat intimidating colleague. Frank Adams and Christopher Zeemen (see Chapter 13) were contemporaries who also influenced me, and I became good friends with Hans Liebeck and Douglas Munn (see Chapter 3), who were close research-student colleagues at St John's.

Yet, the colleague who had the greatest influence on the development of my research during this period, and for many years later, was the cosmologist Dennis Sciama. Even before studying at Cambridge, I had a chance meeting with him during a visit to my brother Oliver in Cambridge in which I raised a question that had worried me in relation to the inspiring series of BBC radio talks given by Fred Hoyle in 1951 (The nature of the universe. A series of broadcast lectures). On the strength of this encounter (at which I used a space-time diagram with tilted light cones to make my point), Dennis had evidently formed the opinion that it would be worth while to interest me more in physics, even trying, later, to persuade me to switch my topic of research to cosmology. I was then too committed to my many pure-mathematical activities, but I learnt a great deal of physics from Dennis, and of the excitements involved in striving to uncover secrets of the universe. Another influential colleague was Felix Pirani, from whom I learnt much about developments in general relativity.

Pivotal were the superb lecture courses of Bondi, on general relativity, and Dirac, on quantum mechanics. On the pure-mathematical side, I recall particularly an elegant information-packed algebra course by Philip Hall. Lectures by Steen on mathematical logic turned out to influence me strongly towards my later viewpoint concerning the relation between computation and conscious understanding (e.g. Chapters 169, 184, 186).

Degrees

1952 BSc Mathematics (Spec) (1st Cl), University College, London

1957 PhD (Cantab), St John's College, Cambridge

Employment

1956	NRDC London (research on numerical analysis with computers); later consultant
1956–1957	Assistant Lecturer, Bedford College, London
1957–1960	Research Fellow, St John's College, Cambridge
1959–1961	NATO Research Fellow, Princeton, Syracuse, Cornell Universities
1961–1963	Research Associate, King's College, London
1963–1964	Visiting Associate Professor, University of Texas, Austin, Texas
1964–1966	Reader (Applied Mathematics), Birkbeck College, London
1966–1967	Visiting appointments: Yeshiva, Princeton, Cornell, Chicago Universities; lecturer Battelle Inst., Seattle
1967–1973	Professor (Applied Mathematics), Birkbeck College, London
1969–1970	Visiting professor, Yeshiva New York, Princeton, and Cornell Universities
1973–1998	Rouse Ball Professor (Mathematics), Oxford; 1998– Emeritus
1978	Miller Fellow, Univ. of California, Berkeley
1982–1987	Lovett Professor (Mathematics), Rice Univ., Houston, Texas, USA (Part-time)
1987–1993	Distinguished Professor (Physics, Mathematics), Syracuse Univ. USA (Part-time)
1993–	Francis and Helen Pentz Distinguished Professor (Physics, Mathematics; part-time)
1998–2001	Gresham Professor (Geometry), Gresham College, London

Honours Prizes include 1971 Heinemann Prize; 1972 Fellow of Royal Society; 1978 Plenary Lecture, International Congress of Mathematicians, Helsinki; 1985 Royal Medal (Royal Society);1988 Wolf Prize (physics, with Hawking); 1990 Albert Einstein Medal; 1994 knighted (services to science); 1994 Polish Academy of Sciences; 1998 US National Academy of Sciences; 2000 Order of Merit; 2004 de Morgan Medal (London Mathematical Society); 2001 Royal Irish Academy (Science); 2008 Copley Medal (Royal Society).

Marriages

1958	Joan Isabel Wedge; divorced 1980
1988	Vanessa Dee Thomas

Children

1963	Christopher Shaun
1964	Toby Nicholas
1966	Eric Alexander
2000	Maxwell Sebastian

Honorary degrees

1992	University of New Brunswick, Canada
1993	University of Surrey
1994	University of Bath
1995	University of London
1996	University of Glasgow
1996	University of Essex
1997	University of St Andrews
1998	University of Santiniketan, India
1998	Open University
2002	University of Southampton
2004	University of Waterloo, Canada
2005	University of Leuven, Belgium
2005	Warsaw University
2005	University of Athens
2006	University of York

Books

1972	Techniques of Differential Topology in Relativity (NSF and SIAM)
1984 and 1986	Spinors and Space-Time (Volumes 1 and 2; Cambridge University Press; with Rindler)
1989	The Emperor's New Mind (Oxford University Press 1989)
1994	Shadows of the Mind (Oxford University Press)
1996	The Nature of Space and Time (Princeton University Press; with Hawking)
1997	The Large, the Small and the Human Mind (Cambridge University Press)
2004	The Road to Reality (Jonathan Cape)

INTRODUCTION

Among the new developments that occurred during the period covered by this Volume was the introduction of a particular notion of 'quasi-local mass-momentum and angular momentum'. This was the topic of my Royal Society paper that constitutes Chapter 132, for which the *Twistor Newsletter* article (TN13) in Chapter 130 provides an outline. For this purpose, a new notion of twistor is introduced that applies to a general spacelike 2-surface $\mathcal{S}$ which has the topology of a sphere S^2. Such a *2-surface twistor*—sometimes referred to as a 'superficial twistor'—is obtained by writing down the twistor equation $\nabla^{(A}_{\ A'}\ \omega^{B)} = 0$ and then selecting only those components of this equation for which the derivative acts purely tangentially to $\mathcal{S}$. This enables us to avoid the consistency conditions that forbid the twistor equation having non-trivial solutions in general curved spacetimes. We obtain two complex equations for two complex quantities, and the Atiyah–Singer theorem can be used to tell us that, in the general case, there is a complex 4-dimensional vector space of solutions, this being the 2-surface twistor space $\mathbb{T}(\mathcal{S})$ assigned to $\mathcal{S}$. These 2-surface twistors act as (spin-$\frac{1}{2}$) helicity-lowering operators, and by combining a pair of them with the Weyl curvature tensor, its spin-2 nature is reduced to spin 1, like electromagnetism, and a Gauss-type charge integral provides a suggested measure of the gravitational source surrounded by $\mathcal{S}$. The different choices of 2-surface twistors are intended to provide the needed 10 components of energy-momentum and angular momentum. Many encouraging results were initially obtained by various students and colleagues—particularly Paul Tod—but then difficulties and anomalies began to emerge, which were not all removed by natural-looking modifications of the definition, indicating that our understanding of what is going on lacked some essential insights—a situation that remains with us today.

An extensive paper with Michael Eastwood and Ronny Wells gave a thorough account of the relation between twistor cohomology and massless fields (Chapter 129). Toby Bailey found how to describe charged sources in twistor theory in terms of relative cohomology and bundles over non-Hausdorff twistor spaces (Chapters 134 and 135), following an earlier observation that George Sparling and I had made concerning the twistor description of the Coulomb field (Chapter 108). I subsequently explored a relation between relative 2nd cohomology and non-Hausdorffness later (Chapters 162 and 169).

Other issues began to take my serious attention during this period, the most important being my increasing conviction that the puzzling issue of quantum measurement will have to be resolved by the appropriate unification of quantum mechanics with general relativity, where this union must involve an actual change in the rules of quantum mechanics, not only in space-time structure. In Chapter 121, I used an argument employing 'Hawking's box' to argue that quantum state reduction should indeed be a gravitational effect, for consistency with the information loss in black holes. I proposed a specific criterion for when a gravitationally induced state-reduction ought to take place (Chapter 154—though later I became convinced that a different proposal is more appropriate). This Volume also contains my first excursions into my proposals that physical basis of consciousness might be related to the issue of quantum state reduction (Chapters 170, 185, and 187).

ACKNOWLEDGEMENTS

Oxford University Press has attempted to trace the copyright holder of all the papers included in these volumes and apologizes to copyright holders if permission to publish in this form has not been obtained.

The author and publishers thank the following individuals for their permissions:

Acta Physics Polonica B **280**
Adam Hilger Ltd **147**
American Chemical Society **206**
American Institute of Physics
American Journal of Physics **19, 24**
Journal of Maths and Physics **15, 26, 27, 29, 39, 41, 58, 84**
Errata **15**
American Mathematical Society **30, 180**
Bulletin of the American Mathematical Society **144**
Contemporary Mathematics **179**
American Physical Society
Physical Review Letters **16, 20, 23, 302**
Reviews of Modern Physics **21**
Annals of Human Genetics **100**
Asian Journal of Mathematics **281**
Astronomische Gesellschaft **295**

Benjamin **36**
Bibliopolis **175**
Blackwell **80, 166, 186, 215**
Annals of the New York Academy of Sciences **60, 190, 294**
British Broadcasting Company **59**
British Journal of Psychology **9**

Cambridge Entomological Club **243**
Cambridge University Press **6, 7, 46, 104, 138, 157, 172, 173, 176, 195, 218, 220, 227, 230, 240, 290, 303, 304**
Behavioural and Brain Sciences **193, 194, 224**
Proceedings of the Cambridge Philosophical Society **3, 4, 5, 32, 42, 50, 126**
Chapman and Hall/CRC Press **107, 124, 125, 127, 128, 130, 136, 139, 142, 143, 146, 149, 150, 154, 155, 156, 160, 161, 164, 170, 171, 178, 189, 191, 198, 201, 203, 209, 229, 234, 293**
Connected Globe **257**
Consigle Nazionale Della Richerche **145**
Cornell University Press **33**

Éditions du Centre National de la Recherche Scientifique **14**
Elsevier **12, 45, 52, 68, 85, 86, 92, 93, 123, 151, 167, 185, 187, 226, 279**
Escher, M.C. images: The M.C. Escher Company B.V. **218, 249, 306**
Eureka **1, 2**

Folio Society **284**
Freeman **56**

Granta **297**
Gresham College **277**

Halsted Press **148**
Hanser **66, 67**

Imperial College Press **264**
Imprint Academic **235, 245**
Indian Academy of Sciences **244**
Indiana University Press **25**
Institute of Electrical and Electronics Engineers **196**
Institute of Physics **217**
Classical and Quantum Gravity **253, 254, 274, 282**
European Journal of Physics **258**
Journal of Physics A: Mathematical and General **95**
International Astronomical Union
International Astronomical Union Symposium Proceedings Series No. 13, Cosmology **72**

International Atomic Energy Agency
Contemporary Physics **38**
International Centre for Theoretical Physics **53, 68**
International Mathematical Union **286**
International University Bremen **305**
Inter-University Centre for Astronomy and Physics **265**
The Italian Astronomical Society
Memorie della Società Astronomica Italiana **200**

Longman **78, 79, 108, 116, 117, 134, 135, 141, 177, 188**

The Massachusetts Institute of Technology Press **228, 246**

National Institute of Sciences of Ukraine **48**
Nature **8, 44, 49**
Nature (Physical Science) **47**
New Scientist **10, 101**
New York Academy of Sciences **294**
Oxford University Press **55, 69, 121, 122, 165, 178, 184, 197, 202, 208, 216, 237, 270, 272, 276, 291, 299**

Pearson **109**
Penguin **273**
Penrose's Unpublished Papers **31**
Pergamon Press **17**
Pitman **73, 75, 76, 77, 82, 87, 89, 90, 94, 102, 103, 110**
Pontificial University **296**
Princeton University Press **28, 268**
Annals of Math **13, 131**
Proceedings of the International Congress of Mathematicians **115**

Rivista del Nuovo Cimento **40**
Routledge **169**
Royal Astronomical Society **183**
The Royal Society
Proceedings of the Royal Society Series A **22, 35, 97, 105, 106, 132**
Philosophical Transactions of the Royal Society of London. Series A **275, 283**

The Sciences **99**
Scientific American **51**
Simon and Schuster **242**
Springer **61, 63, 65, 83, 91, 98, 111, 133, 168, 204, 207, 231, 255, 263, 278, 292, 298, 300**
Artificial Intelligence Review **205**
Communications in Mathematical Physics **43, 54, 129**
Cosmology and Gravitation **120**
Foundations of Physics **140**
General Relativity and Gravitation **70, 71, 113, 114, 239, 251**
Higher-Order and Symbolic Computation **287**
International Journal of Theoretical Physics **37, 56**
Journal of Statistical Physics **233**
Lecture Notes in Physics **112, 152**
The Mathematical Intelligencer **285**
Taylor and Francis **18**
International Studies in the Philosophy of Science **256**
Surveys in High Energy Physics **119**
Trinity College, Dublin **219**

World Scientific **162, 163, 225, 236, 241, 252, 289, 301**

Universal Academy Press **266**
University of Chicago Press **96, 269**
University of Chittagong **158**
University College London **267**

The University of Oxford (Mathematics Department)
Twistor Newsletter **73, 74, 75, 76, 77, 78, 79, 81, 82, 88, 89, 90, 94, 102, 103, 107, 108, 116, 117, 118, 124, 125, 127, 128, 130, 134, 135, 136, 137, 139, 141, 142, 143, 146, 149, 150, 153, 154, 155, 156, 157, 160, 161, 164, 170, 171, 174, 181, 182, 188, 189, 191, 192, 198, 199, 201, 203, 209, 210, 211, 212, 213, 214, 221, 222, 223, 229, 232, 238, 248, 290, 250, 259, 260, 261, 262, 271, 288**

University of Utah Press **247**

CONTENTS
VOLUME 4

	Introduction	*xxv*
	1980s	
121	Time-asymmetry and quantum gravity	1
122	Some remarks on twistors and curved-space quantization	31
123	On Schwarzschild causality—a problem for "Lorentz covariant" general relativity	47
124	Concerning a Fourier contour integral	61
125	A new angle on the googly graviton	65
126	A quadratic mapping with invariant cubic curve	73
127	Twistors for cosmological models	91
128	Physical left-right symmetry and googlies	97
129	Cohomology and massless fields	105
130	A theory of 2-surface ('superficial') twistors	153
131	Some unsolved problems in classical general relativity	161
132	Quasi-local mass and angular momentum in general relativity	201
133	Time-asymmetry, cosmological uniformity and space-time singularities	213
134	Real-space topology for advanced/retarded twistor functions	217
135	Twistor blisters and relative cohomology	221
136	On the geometry of googly maps	225
137	A puzzle for twistorians	229
138	Some remarks on gravity and quantum mechanics	231
139	General-relativistic kinematics?	241
140	Spinors and torsion in general relativity	247
141	Non-locality in physics: A new mathematical viewpoint	263
142	On closed-set coverings for PT^+—A correction	269
143	Remarks on curved-space, twistor theory and googlies	271
144	A note on Sparling's 3-form	277
145	Physical space-time and nonrealizable CR-structures	281
146	Quasi-local mass and angular momentum	305
147	Relative cohomology, googlies and deformations of $\mathbb{I}$	309

148	Donaldson's moduli space: A 'model' for quantum gravity?	313
149	Mass in general relativity	319
150	'New improved' quasi-local mass, and the Schwarzschild solution	331
151	On Bryant's condition for holomorphic curves in CR space	339
152	Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory	343
153	Mass and angular momentum at the quasi-local level in general relativity	377
154	Objective state-vector reduction?	387
155	Embedding 2-surfaces in CM	393
156	A suggested further modification to the quasi-local formula	397
157	Higher-dimensional two-surface twistors	399
158	Numerical relativity by power series	403
159	Elements of twistor theory	409
160	Energy and its definition in general relativity	429
161	Two-surface twistors and hypersurface twistors	441
162	Exponentiating a relative H^2	445
163	Gravitational field mass	449
164	Gravity and quantum theory	467
165	Twistors and the time-irreversability of state-vector reduction	473
166	Gravity and state vector reduction	477
167	Big bangs, black holes and 'time's arrow'	497
168	Escher and the visual representation of mathematical ideas	525
169	Hermann Weyl, space-time and conformal geometry	541
170	Quantum physics and conscious thought	573
171	Local twistor transport at $\mathscr{I}^+$: An approach to the googly	601
172	An approach to a coordinate-free calculus at $\mathbb{I}$	609
173	Twistors in general relativity	613
174	Newton, quantum theory and reality	633
175	On Qadir's intertwining operators	667
176	On the origins of twistor theory	671
177	New directions in quantum gravity	693
178	Pretzel twistor spaces	703
179	Twistors and state-vector reduction	713
180	Aspects of quasi-local angular momentum	717
181	Fundamental asymmetry in physical laws	727
182	Topological QFT and twistors: Holomorphic linking	741

183	Holomorphic linking: Postscript	745
184	Singularities and big-bang cosmology	747
185	On the physics and mathematics of thought	751
186	A twistor conformal field theory for four space-time dimensions	785
187	Minds, machines and mathematics	791
188	Tilings and quasi-crystals: A non-local growth problem?	811
189	Non-Hausdorff Riemann surfaces and complex dynamical systems	839
190	Using spinors in general relativity	845
191	A twistorial approach to the full vacuum equations	851

DETAILED CONTENTS
VOLUMES 1–6

VOLUME 1

1950s

1 Morley's trisector theorem 1
2 'Construction problem' and 'Word problem' 3
3 A note on inverse semigroups 5
4 A generalized inverse for matrices 11
5 On best approximate solutions of linear matrix equations 21
6 Tensor methods in algebraic geometry 25
7 Geometrical algebras: A new approach to invariant theory 251
8 A self-reproducing analogue 351
9 Impossible objects: A special type of visual illusion 355
10 Puzzles for Christmas 359
11 The apparent shape of a relativistically moving sphere 365

1960s

12 A spinor approach to general relativity 369
13 Imbedding of manifolds in Euclidean space 403
14 General relativity in spinor form 415
15 An approach to gravitational radiation by a method of spin coefficients 421
16 Asymptotic properties of fields and space-times 437
17 The light cone at infinity 441
18 Conformal treatment of infinity 447
19 Energy conservation as the basis of relativistic mechanics 469
20 Gravitational collapse and space-time singularities 475
21 A remarkable property of plane waves in general relativity 479
22 Zero rest-mass fields including gravitation: Asymptotic behaviour 487
23 10 exact gravitationally conserved quantities 533
24 Energy conservation as the basis of relativistic mechanics II 537
25 General-relativistic energy flux and elementary optics 541
26 Note on the Bondi-Metzner-Sachs group 559
27 Some new gravitationally conserved quantities 569
28 An analysis of the structure of space-time 579
29 Twistor algebra 731
30 Conserved quantities and conformal structure in general relativity 755
31 Theory of quantized directions 769
32 On the structure of causal spaces 801

VOLUME 2

1960s

33 Cosmological boundary conditions for zero rest-mass fields 1
34 On gravitational collapse and cosmology 15
35 New conservation laws for zero rest-mass fields in asymptotically flat space-time 29
36 Structure of space-time 61
37 Twistor quantization and curved space-time 177
38 On gravitational collapse 217

39 Solutions of the zero-rest-mass equations 231
40 Gravitational collapse: The role of general relativity 235
41 Conserved quantities in the Einstein–Maxwell theory 261

1970s

42 The singularities of gravitational collapse and cosmology 267
43 On quadratic first integrals of the geodesic equations for type {22} spacetimes 289
44 Scattering of two impulsive gravitational plane waves 301
45 Applications of negative dimensional tensors 307
46 Angular momentum: An approach to combinatorial space-time 333
47 Extraction of rotational energy from a black hole 367
48 Spinor classification of energy tensors 375
49 Black holes and gravitational theory 389
50 Ideal points in space-time 395
51 Black holes 419
52 Twistor theory: An approach to the quantisation of fields and space-time 429
53 a) 'Radiation, energy and the BMS group'; b) 'Some recent developments'; c) 'N-P constants' 505
54 On a quadratic first integral for the charged particle orbits in the charged Kerr solution 513
55 The geometry of impulsive gravitational waves 521
56 On the nature of quantum geometry 537
57 Internal instability in a Reissner–Nordström black hole 583
58 A space-time calculus based on pairs of null directions 599
59 Black holes 609
60 Naked singularities 635
61 Relativistic symmetry groups 647
62 Gravitational collapse 707
63 Gravitational collapse 729
64 The role of aesthetics in pure and applied mathematical research 741
65 Singularities in cosmology 749
66 Twistors and particles: An outline 761
67 The non-linear graviton 779
68 Classical gravity: A status report 791
69 Twistor theory: Its aims and achievements 795

VOLUME 3

1970s

70 Nonlinear gravitons and curved twistor theory 1
71 The nonlinear graviton 25
72 The space-time singularities of cosmology and in black holes 33
73 The twistor quadrille 45
74 The snail contour 53
75 The universal bracket factor 55
76 Some elementary twistor integrals with boundary 59
77 Twistor functions and sheaf cohomology 63
78 How to grow (or at least half grow) new twistors from old 77
79 Some remarks on the two twistor description of electromagnetic and linearized gravitational potentials 81
80 The cosmic censorship hypothesis 83
81 A finite covering for PT^+ 87
82 The back-handed photon (or, to a cricketer, the googly photon!) 89
83 Any space-time has a plane wave as a limit 99
84 The Kähler structure of asymptotic twistor space 105
85 The twistor programme 113
86 Space-time singularities 127
87 The twisted camel (or how to get through the eye of a singularity). Includes: more on the twisted camel 137
88 Remarks concerning the ambidextrous photon 145
89 Massless fields and sheaf cohomology 149
90 Note on the n-twistor internal symmetry group 159
91 Aspects of general relativity 163
92 Is nature complex? 203
93 Gravitational collapse: A review 213
94 Local H^1's and propagation 231

95 Extended Regge trajectories 237
96 Singularities of spacetime 243
97 The metric and curvature properties of $\mathcal{H}$-space 271
98 The geometry of the universe 297
99 Einstein's vision and the mathematics of the natural world 341
100 The topology of ridge systems 347
101 Twisting round space-time 359
102 A googly graviton? 365
103 Twistors as helicity raising operators 375
104 Singularities and time-asymmetry 381
105 General introduction 441
106 Global general relativity 445
107 Remarks on the Sparling and Eguchi-Hanson (googly?) gravitons 449
108 The anti-self-dual Coulomb field's non-Hausdorff twistor space 459
109 On the twistor descriptions of massless fields 461
110 Combinatorial quantum theory and quantized directions 499
111 Pentaplexity: A class of nonperiodic tilings of the plane 507

1980s

112 Recent advances in global general relativity: A brief survey 515
113 Twistors for flat and curved space-time 527
114 Null hypersurface initial data for classical fields of arbitrary spin and for general relativity 575
115 The complex geometry of the natural world 617
116 On the evaluation of twistor cohomology classes 625
117 Three channels for the box diagram 629
118 Puzzle: A sequence for twistorians 637
119 A brief introduction to twistors 639
120 A brief outline of twistor theory 663

VOLUME 4

1980s

121 Time-asymmetry and quantum gravity 1
122 Some remarks on twistors and curved-space quantization 31
123 On Schwarzschild causality—a problem for "Lorentz covariant" general relativity 47
124 Concerning a Fourier contour integral 61
125 A new angle on the googly graviton 65
126 A quadratic mapping with invariant cubic curve 73
127 Twistors for cosmological models 91
128 Physical left-right symmetry and googlies 97
129 Cohomology and massless fields 105
130 A theory of 2-surface ('superficial') twistors 153
131 Some unsolved problems in classical general relativity 161
132 Quasi-local mass and angular momentum in general relativity 201
133 Time-asymmetry, cosmological uniformity and space-time singularities 213
134 Real-space topology for advanced/retarded twistor functions 217
135 Twistor blisters and relative cohomology 221
136 On the geometry of googly maps 225
137 A puzzle for twistorians 229
138 Some remarks on gravity and quantum mechanics 231
139 General-relativistic kinematics? 241
140 Spinors and torsion in general relativity 247
141 Non-locality in physics: A new mathematical viewpoint 263
142 On closed-set coverings for PT^+—A correction 269
143 Remarks on curved-space, twistor theory and googlies 271
144 A note on Sparling's 3-form 277
145 Physical space-time and nonrealizable CR-structures 281
146 Quasi-local mass and angular momentum 305
147 Relative cohomology, googlies and deformations of $\mathbb{I}$ 309
148 Donaldson's moduli space: A 'model' for quantum gravity? 313
149 Mass in general relativity 319
150 'New improved' quasi-local mass, and the Schwarzschild solution 331
151 On Bryant's condition for holomorphic curves in CR space 339

152 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory 343
153 Mass and angular momentum at the quasi-local level in general relativity 377
154 Objective state-vector reduction? 387
155 Embedding 2-surfaces in CM 393
156 A suggested further modification to the quasi-local formula 397
157 Higher-dimensional two-surface twistors 399
158 Numerical relativity by power series 403
159 Elements of twistor theory 409
160 Energy and its definition in general relativity 429
161 Two-surface twistors and hypersurface twistors 441
162 Exponentiating a relative H^2 445
163 Gravitational field mass 449
164 Gravity and quantum theory 467
165 Twistors and the time-irreversability of state-vector reduction 473
166 Gravity and state vector reduction 477
167 Big bangs, black holes and 'time's arrow' 497
168 Escher and the visual representation of mathematical ideas 525
169 Hermann Weyl, space-time and conformal geometry 541
170 Quantum physics and conscious thought 573
171 Local twistor transport at $\mathscr{I}^+$: An approach to the googly 601
172 An approach to a coordinate-free calculus at $\mathbb{I}$ 609
173 Twistors in general relativity 613
174 Newton, quantum theory and reality 633
175 On Qadir's intertwining operators 667
176 On the origins of twistor theory 671
177 New directions in quantum gravity 693
178 Pretzel twistor spaces 703
179 Twistors and state-vector reduction 713
180 Aspects of quasi-local angular momentum 717
181 Fundamental asymmetry in physical laws 727
182 Topological QFT and twistors: Holomorphic linking 741
183 Holomorphic linking: Postscript 745
184 Singularities and big-bang cosmology 747
185 On the physics and mathematics of thought 751
186 A twistor conformal field theory for four space-time dimensions 785
187 Minds, machines and mathematics 791
188 Tilings and quasi-crystals: A non-local growth problem? 811
189 Non-Hausdorff Riemann surfaces and complex dynamical systems 839
190 Using spinors in general relativity 845
191 A twistorial approach to the full vacuum equations 851

VOLUME 5

1990s

192 Difficulties with inflationary cosmology 1
193 Light rays near i^0: A new mass-positivity theorem 19
194 Extended Regge trajectories: 12 years after 27
195 Précis of *The Emperor's New Mind: Concerning computers, minds and the laws of physics* 29
196 Author's response – the nonalgorithmic mind 43
197 Twistor theory after 25 years – its physical status and prospects 59
198 Twistors, particles, strings and links 89
199 Mass positivity from focussing and the structure of space-like infinity 101
200 Twistor theory for vacuum space-time: A new approach 107
201 What does the big bang tell us about quantum gravity? 113
202 Twistors as spin 3/2 charges 123
203 Twistors as charges for spin 3/2 in vacuum 133
204 The mass of the classical vacuum 143
205 Twistors as spin 3/2 charges continued: SL(3,C) bundles 151
206 Response to Tony Dodd's 'Godel, Penrose and the possibility of AI' 159
207 The biggest enigma 161
208 $\mathcal{H}$-space and twistors 167
209 'Penrose and mathematical ability' 189

210 The most general {2,2} self-dual vacuum: A googly approach 193
211 On impulsive gravitational waves 199
212 Kinking and causality 205
213 Puzzle page 213
214 Solution to puzzle in TN34 215
215 Gravitational collapse of the wave function: New thoughts 221
216 Setting the scene: The claim and the issues 225
217 The modern physicist's view of nature 259
218 Gravity and quantum mechanics 311
219 Foreword to *Mr Tompkins in paperback* 323
220 Mathematics, physics, computers and conscious understanding 327
221 An example of indeterminacy in the time-development of "already unified field theory": A collision between electromagnetic plane waves 339
222 Orthogonality of general spin states 349
223 A simple observation concerning {2 2} vacuums 355
224 The orthographic image of a regular tetrahedron 357
225 Author's response – An Emperor still without mind 361
226 Square-root voting 369
227 On Bell non-locality without probabilities: More curious geometry 375
228 Quantum non-locality and complex reality 401
229 On the cohomology of impossible figures 415
230 Spin 3/2 fields and local twistors 419
231 Mathematical intelligence 427
232 Is conscious awareness consistent with space-time descriptions? 459
233 A twistor-topological approach to the Einstein equations 475
234 On the second law of thermodynamics 485
235 Twistors and the Einstein equations 491
236 What 'gaps'? Reply to Grush and Churchland 505
237 Non-locality and objectivity in quantum state reduction 521
238 Must mathematical physics be reductionist? 531
239 Concerning space-time points for spin 3/2 twistor spaces 547
240 In {2 2} vacuums, the null-datum on null rays is exactly r^{-5} 553
241 Why new physics is needed to understand the mind 557
242 The role of gravity in quantum state reduction 573
243 Consciousness involves noncomputable ingredients 593
244 Beyond the doubting of a shadow: A reply to commentaries on *Shadows of the Mind* 603
245 Chandrasekhar, black holes, and singularities 641
246 Conscious events as orchestrated space-time selections 661
247 Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness? 681
248 Space-time and cosmology 711
249 New ideas for the googly graviton 757
250 Incidence between complex null rays 771
251 Two puzzles 777
252 On gravity's role in quantum state reduction 779
253 Twistor theory, the Einstein equations, and quantum mechanics 801

VOLUME 6

1990s

254 Andrzej Mariusz Trautman 1
255 Twistor geometry of light rays 11
256 Remarks on tiling: Details of a $(1+\varepsilon+\varepsilon^2)$-aperiodic set 37
257 On understanding understanding 69
258 Mathematics: Pattern, precision and profundity 85
259 The mathematics of the electron's spin 109
260 Five things you can do with a surface in PT 115
261 Solutions to puzzles in TN41 121
262 Corrections to missprints in solutions to TN41 puzzles 125
263 Googly maps as forms 127
264 The need for a non-computational extension of quantum action in the brain 133
265 Chandrasekhar and the end of time 151
266 Science and the mind 161

267 Quantum state-reduction as a gravitational phenomenon; theory and experiment 165
268 Lionel Penrose: Colleague and father 169
269 Foreword to *Einstein's miraculous year: Five papers that changed the face of physics* 175
270 The question of cosmic censorship 185
271 Afterword to *The geometric universe: Science, geometry, and the work of Roger Penrose* 203
272 The twistor space for a strongly asymptotically flat vacuum 213
273 The complexity of our singular universe 223
274 Can a computer understand? 251
275 Spherically-symmetric solutions of the Schrödinger–Newton equations 279
276 Quantum computation, entanglement and state reduction 291
277 Foreword to *Ludwig Boltzmann: The man who trusted atoms* 305
278 Mathematics in the arts 309
279 Some remarks on twistor theory 323
280 The central programme of twistor theory 339
281 Lecture in honour of Leopold Infeld (extended outline only): Spinors in general relativity 371
282 On Ricci-flat twistor theory 381
283 Twistor theory and the Einstein vacuum 389
284 Introduction: Science at the turn of the millennium. 'Science into the next millennium: Young scientists give their visions of the future, part I.' 409

2000s

285 Introduction. *What is life?* 417
286 Drawing with complex numbers 433
287 Mathematical physics of the 20th and 21st centuries 441
288 Reminiscences of Christopher Strachey 459
289 On extracting the googly information 463
290 Wavefunction collapse as a real gravitational effect 489
291 On Bell non-locality without probabilities: Some curious geometry 509
292 The heritage of Pythagoras: Nineteen to the dozen 537
293 Twistors and general relativity 551
294 Towards a twistor description of general space-times: Introductory comments 579
295 Consciousness, the brain, and spacetime geometry: An addendum 597
296 The Schwarzschild singularity 605
297 The physics of the small and large: What is the bridge between them? 621
298 The rediscovery of gravity. The Einstein equation of general relativity 629
299 John Bell, state reduction, and quanglement 663
300 Consciousness, computation, and the Chinese room 677
301 Twistor geometry of conformal infinity 703
302 Gravitational collapse of the wavefunction: An experimentally testable proposal 713
303 Towards quantum superpositions of a mirror 719
304 On the instability of extra space dimensions 725
305 The problem of spacetime singularities: Implications for quantum gravity? 743
306 The remarkable role of mathematics in the physical universe 767

121
TIME-ASYMMETRY AND QUANTUM GRAVITY

Quantum gravity 2: A second Oxford symposium. Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. Clarendon Press, Oxford, 1981

As in Chapter 104, I again express my disbelief in Hawking's identification between black holes and white holes. Moreover, the fundamental link between the Weyl curvature hypothesis (or something of this nature) and the second law of thermodynamics is again stressed. But this is the first presentation of my argument using 'Hawking's box' to demonstrate a link between the quantum measurement process and the 'information-loss' issue with regard to black holes. (The reader should perhaps be warned that my labelling of unitary evolution as 'Rule 1' and quantum state reduction as 'Rule 2' appears to follow the opposite convention from that adopted by von Neuman.)

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

11
TIME-ASYMMETRY AND QUANTUM GRAVITY

R. Penrose
Mathematical Institute, University of Oxford

At our previous Oxford Quantum Gravity Conference, held in 1974, Stephen Hawking put forward his remarkable theory of black-hole evaporation (Hawking 1975*a*,*b*). It seemed clear at the time that these ideas were destined to play a key role in any future development in which quantum field theory and general relativity were to be successfully united. Two years later, in 1976, we held a conference which was specifically devoted to the furtherance of these ideas. But by then, there had begun to emerge a certain divergence of opinion between Hawking and myself – something which we still have not found ourselves able to resolve. A point on which we are certainly agreed, however, is that this conflict has proved immensely stimulating to us both. My own line of argument has been driven, in part under the sting of Hawking's penetrating criticisms, into some dangerous but fascinating territory. I hope my audience will bear with me if I may seem to be drawing some insufficiently supported, and perhaps somewhat radical conclusions concerning the nature of quantum-mechanical observations. But to me, though they are yet very rudimentary, these conclusions seem to be pointing in a necessary direction for future physics. I hope, also, that I may be forgiven if I do not adequately describe Hawking's own viewpoint – and undoubtedly I shall do it considerably less than justice. But, however inadequately, I shall need to give some indications of it in order to present the reasons behind my own very different point of view more clearly.

1. HOW SPECIAL WAS THE BIG BANG?

Let us first recall the Bekenstein–Hawking formula (Bekenstein 1973, 1974; Hawking 1975*a*) for the entropy of a stationary

black hole with a horizon of surface area A:

$$S_{bh} = \frac{kAc^3}{4\hbar G}$$

$$= \tfrac{1}{4}A \quad (\text{with } k = c = \hbar = G = 1).$$

(k = Boltzmann's constant; c = velocity of light; $2\pi\hbar$ = Planck's constant; G = Newton's gravitational constant.) One way in which this formula can be used is to get an estimate of how 'special' the big bang was (Penrose 1979). In order to fix ideas, let us consider a closed universe containing $\sim 10^{80}$ baryons. The standard figure, assuming our universe to be closed, for the observed entropy per baryon (essentially, the number of photons per baryon and taking $k = 1$) is $\sim 10^{8}$. This does not take into account the possibility of an appreciable contribution to the total entropy arising from the presence of black holes. We shall see how black holes could affect the discussion in a moment. Another problem with this figure is the strong possibility that in the very early stages there may have been an appreciable non-conservation of baryon number. Such ideas have become very fashionable recently, particularly in connection with the 'grand unified theories'. For this reason it is sometimes useful to talk in terms of the *total* entropy of the universe, the 'observed' figure therefore being 10^{88}.

Now consider what the Bekenstein–Hawking formula tells us should be the entropy per baryon for a single solar mass black hole. The figure is $\sim 10^{20}$. This is some twelve orders of magnitude larger than the 'observed' figure. Thus if our (assumed closed) universe had consisted, at our epoch, entirely of solar-mass black holes, the *total* entropy would have been about twelve orders of magnitude larger than that which is apparently observed. Of course, our universe is not so constituted! But these figures indicate something about the 'observed' entropy per baryon, namely that in a gravitational context it must be regarded as an exceedingly small value.

Before the discovery of the Bekenstein–Hawking formula it was perhaps possible to maintain that this 'observed' figure represented a surprisingly high value for the entropy per baryon in the universe. Indeed, the programme of chaotic

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

cosmology was proposed (Misner 1968, 1969; Zel'dovich 1972) according to which an initially chaotic geometry was supposed to be responsible for the 'high' present entropy figure, dissipation effects serving to iron out most of the irregularities in the initial geometry and transferring whatever entropy was contained in these irregularities into the black-body radiation. We shall see in a moment how absurdly huge these constraints on the initial geometry would have had to have been in order not to have grossly exceeded the present observed figure for the entropy (cf. Penrose 1974, 1979 and also Barrow and Matzner 1977 for a more specific detailed argument). But even without performing any numerical calculations, one may have strong reservations about the chaotic cosmology idea if, in some sense, the initial chaos was supposed to be at a maximum. For then the universe would have had to have started in a 'maximum entropy state', dissipation would not occur, and there would be no second law of thermodynamics.* The entropy of the universe would have already reached its maximum and there would be no further way in which the universe could 'relax'.

Though the phenomenon of the expansion of the universe serves to obscure this discussion somewhat, it is *not* true that the mere fact of expansion alone could lead to an entropy increase and itself provide us with our second law of thermodynamics. (I have gone into this question at some length elsewhere (Penrose 1979) and have no wish to repeat the argument here, though some further comments are given in the next section.) In is necessary, in fact, that the structure of the big bang singularity was severely constrained in order that thermodynamics as we know it should have arisen.

In order to form an impression of just *how* constrained it

---

* It should be made clear here that the second law *is* something in need of explanation, in our universe. For while it is not surprising that a system in a given state of low entropy should normally proceed, in the future direction, to a state of higher entropy, it is a remarkable fact of our world that this argument does *not* apply in the *past* direction. As we work backwards to earlier and earlier epochs we find that the entropy goes down and down. For the second law to operate as we know it, and to have operated as we know it at all times in the past, it is necessary that the earliest stages of the universe should have been constrained in some very precise way. The puzzle is: where do these constraints come from?

must have been, let us proceed further in our application of the Bekenstein–Hawking formula. Rather more realistic than galaxies populated entirely with solar mass black holes would be, say, 10^{11} $M_\odot$ galaxies with 10^6 $M_\odot$ black-hole cores. Here the entropy per baryon would be 10^{21} – an order of magnitude higher again than the solar mass black hole case – and the total entropy for a 10^{80} baryon universe entirely constituted in this way would be 10^{101}. (Having large black holes enables us to gain enormously in the entropy and to compensate easily for the fact that now most of the mass is in the form of ordinary stars.) Of course we can gain much more in the total entropy by allowing each entire galaxy to be swallowed by its black hole core. For the resulting $10^{11}M_\odot$ black hole the entropy per baryon is now 10^{31}, and for a 10^{80} baryon universe entirely composed of such objects the total entropy would be 10^{111} – some 23 orders of magnitude higher than the supposedly high 'observed' figure of 10^{88}.

In the final collapsing phase of the universe, irrespective of the particular one of these ways in which it might happen to have been constituted, we expect smaller black holes to unite into larger and larger ones until ultimately a figure of the order of 10^{43} for the entropy per baryon is achieved, giving 10^{123} for the total entropy of the universe. To obtain this figure, we cannot quite apply the Bekenstein–Hawking formula directly since the final state of the universe does not at that stage, at all geometrically resemble the isolated black-hole model to which the Bekenstein–Hawking formula applies. However, the figure seems a reasonable estimate, since only a little before that stage the universe would have consisted essentially of a few black holes more or less separated from one another, the figure of 10^{123} for the total entropy being approached from below as the holes all finally unite together. Thus it would seem that the figure of 10^{123} is a plausible estimate for the maximum entropy state of a universe of this type. (See Table 1.1)

If we accept this figure, and compare it with the original 'observed' figure of 10^{88} for the entropy which, according to the chaotic cosmology view, should have been produced out of initial irregularities of the big bang, then we realize

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

TABLE 1.1

Entropy values for a closed universe with B = 10^{80} *baryons (Boltzmann's constant = 1)*

	S/B	corresponding total *S*
'observed'	10^{8}	10^{88}
1 $M_{\odot}$ b.h.	10^{20}	10^{100}
10^{6} $M_{\odot}$ b.h. in 10^{11} $M_{\odot}$ galaxy	10^{21}	10^{101}
10^{11} $M_{\odot}$ b.h.	10^{31}	10^{111}
collapsed universe	10^{43}	10^{123}

how absurdly tiny this 'observed' figure is in comparison with what it 'might have been'. This provides us with a measure of the degree to which the initial state was special. Indeed, even the figure 10^{88} totally ignores the entropy increase that would have had to have been involved in the dissipation processes responsible for converting the irregular initial geometry into background radiation. According to the chaotic cosmology view, the figure of 10^{88} for the 'observed' total entropy must be considerably higher than the 'actual' value for the entropy in the initial chaotic geometry.

To get a quantitative picture of the degree which the big bang was special imagine a manifold* $\mathscr{W}$ whose points represent the various possible initial configurations of the universe. Imagine the Creator, armed with a pin which is to be placed at one spot in $\mathscr{W}$ thereby determining the state of our actual universe. How accurate must the Creator's aim have been in order that a universe of the 'observed' type could have arisen? Since entropy is, after all, related to a phase-space volume w in $\mathscr{W}$ by the logarithmic formula $S = k \log w$ ($k = 1$), we are led to estimate that the accuracy of the Creator's aim must

* I am using the word 'manifold' loosely here. No requirement is implied that it be a manifold in the technical sense.

have been at least of the order* of $10^{10^{88}}$ parts in $10^{10^{123}}$ (this being the ratio of the volume to be aimed at to the total volume of $\mathcal{W}$) i.e. one part in†

$$10^{(10^{123}-10^{88})} \dot{=} 10^{10^{123}} \quad .$$

Even if the universe had consisted almost entirely of 10^{11} $M_\odot$ black holes, the Creator's aim would have had to have been accurate to effectively the same degree as before since replacing '88' by '111' in this expression will make no difference to it. Without wishing to denigrate the Creator's abilities in this respect, I would insist that it is one of the duties of science to search for physical laws which explain, or at least describe in some coherent way, the nature of the phenomenal accuracy that we so often observe in the workings of the natural world. Moreover, I cannot even recall seeing anything else in physics whose accuracy is known to approach, even remotely, a figure like one part in $10^{10^{123}}$. So we need a *new law of physics* to explain the specialness of the initial state!

At this point I should make clear something which might otherwise be misleading. It may be felt that, in a chaotic cosmology viewpoint the reason for the time-asymmetry of the world has merely to do with the Creator's choice being made

---

* Some peculiarities of operating with such large numbers should be pointed out (see Littlewood 1953). It makes no essential difference, for example, whether one writes

$10^{10^{88}}$, or $e^{10^{88}}$, or even $1000^{10^{88}}$, for that matter, since

$1000^{10^{88}} = 10^{3\times10^{88}} \dot{=} 10^{10^{88.5}}$, while $e^{10^{88}} \dot{=} 10^{0.43\times10^{88}} \dot{=} 10^{10^{87.6}}$. In such expressions, the symbol '$\dot{=}$' must be interpreted suitably. It does not now mean that the ratio of the two sides of the equation is close to unity, but that it is not large enough to imply a substantial change in the uppermost exponent. (Note, that, in this sense, even cubing the number $10^{10^{88}}$ makes no essential difference to it!) The uncertainties in the figure '88' will quite swamp any of these other considerations.

† In an earlier work (Penrose 1979) there is, implicitly, an oversight which amounts to writing the figure erroneously as

$10^{(10^{120}/10^{88})} = 10^{10^{32}}$. Thus the Creator's aim would need to be enormously more accurate even than is suggested in that reference.

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

in the remote past rather than in the remote future. However, if we accept the standard picture of classical mechanics, this is not the case. I am doubtful, also, that it would make any real difference if a conventional quantum mechanical description is adopted instead of a classical one. But let us, for definiteness, suppose that the universe evolves classically according to standard Hamiltonian-type equations. Then the phase-space measure on the manifold $\mathscr{W}$ remains constant as the universe evolves (Liouville's theorem), where we now envisage that each point of $\mathscr{W}$ represents not just the initial conditions, but the entire 4-dimensional universe-history (i.e. space-time) that evolves from this initial state. It makes no difference to the remarkableness of the Creator's selection, whether this is in choosing initial or final conditions, or simply selecting one entire space-time history out of all the vast collection of alternative possibilities. The figure of one part in $10^{10^{123}}$, which represents the improbability of the Creator's selection, still applies whether the choice is made at the beginning or at the end, or as a selection of an entire universe history, since it simply represents a ratio of phase-space volumes.

In this, I am taking that figure as a measure of the improbability of our actual universe (assumed closed, for the purposes of this argument), if we *assume* that the Creator's choice was made entirely at random – for this is the essential hypothesis of the chaotic cosmology viewpoint. I think that it should be clear from the above discussion that this viewpoint cannot actually be valid, in our world, if we assume an unconstrained initial state. The 'pure' chaotic cosmology viewpoint has, indeed, fallen out of favour in recent years. But even if we adopt a viewpoint of 'partially constrained chaos' according to which *some* geometrical constraint is placed on the initial singularity, we still have the problem of finding how such a constraint can produce something so absurdly close to uniformity while not yielding *exact* uniformity in the initial geometry. In my view, this is a clear indication that chaos in the initial geometry was *not* what was responsible for the present entropy of the universe. The 'observed' figure of 10^8 for the entropy per baryon of the

present universe must have its origin in some part of physics other than gravitation theory. The most likely candidate would seem to be particle physics, but I shall make no attempt to pursue the matter further here.

2. IS A TIME-SYMMETRIC UNIVERSE TENABLE?

In Fig. 2.1 is a crude qualitative picture which contrasts the 'universe as we know it' (assumed closed, for our expressed purpose of convenience only) with a 'more probable universe' having a chaotic big bang. The big crunch is, in each case, a high entropy ($\sim 10^{123}$) complicated unconstrained final singularity, while, for the left-hand picture the big bang is a low entropy ($\leqslant 10^{88}$) Friedmann-like, highly constrained initial singularity, the right-hand picture having an unconstrained 'much more probable' big bang. The diagrams are suggestive only, and not drawn strictly according to the rules of standard conformal diagrams whereby null cones would be drawn as sloping at 45°. The 'stalactites' represent the singularities of black holes, while the 'stalagmites' of the second picture represent the singularities of white holes.

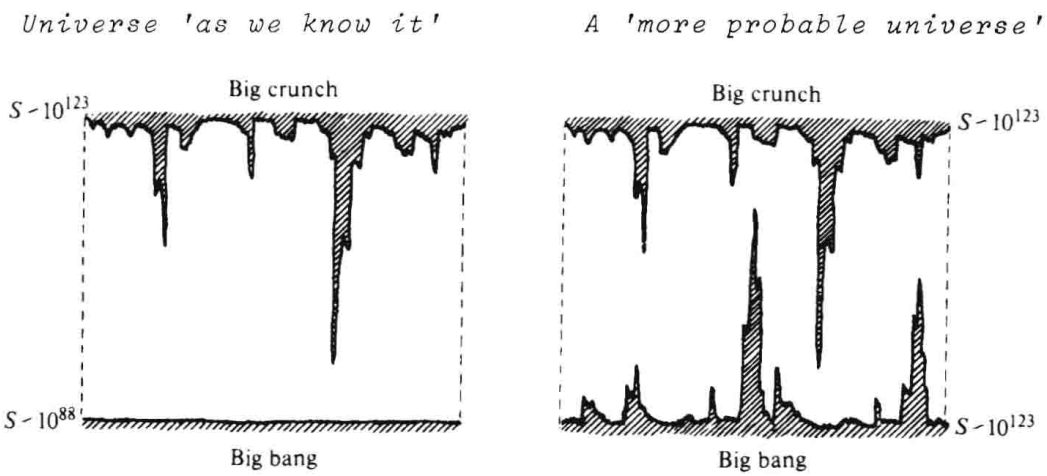


FIG. 2.1. If there were no severe constraint placed on the big bang, as there is in the diagram on the left, we would be driven to the picture of a white-hole-riddled universe on the right. (Not drawn strictly according to conformal diagram conventions.)

 Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

Note that it is primordial white holes, rather than primordial black holes that seem to occur most naturally in a chaotic big bang. For the 'chaos' is simply the time-reverse of that which we expect to arise naturally in the generic collapse of a universe such as our own (assumed closed). Primordial black holes could occur too, but they represent a more complicated geometrical situation (see Fig. 2.2).

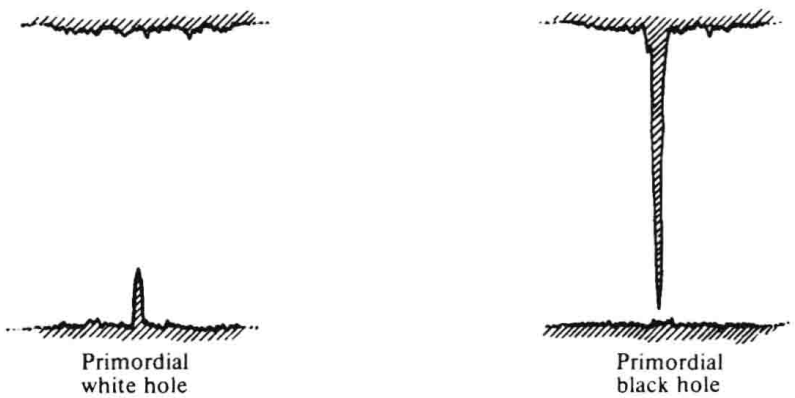


FIG. 2.2. The geometry of the singularity structure for primordial white and black holes.

It should be pointed out, also, that the various arguments in the literature which indicate that white holes ought, in some sense to be violently *unstable* (Eardley 1974; Zel'dovich 1974; Ramaswamy and Wald 1980) depend upon assuming that there is an otherwise 'normal' thermodynamic time-arrow and that the white hole can be treated as isolated. Although not directly applicable in the present context, these arguments seem to imply that a white-hole singularity which can be well separated from the remainder of the big-bang singularity is a highly 'improbable' type of initial singularity, requiring very precisely organised correlations in the incoming fields (though even so, it is not as 'improbable' as the Friedmann-type singularity itself, among all unconstrained big bangs). The reason that the time-reverse of this – namely an isolated *black* hole – can occur as a common feature of a 'normal' universe is that correlations in the outgoing fields are to be *expected* owing to the fact that such a 'normal' universe has a low-entropy Friedmann-type big bang (and hence a second law of thermodynamics), this low entropy being entirely

reflected in such correlations when the final state is reached. In time-reversed form, the correlations needed to keep an isolated white hole in existence would require a low-entropy *final* state. I certainly do *not* want to assume a low-entropy final state, but this should not restrict the freedom in the choice of structure for the initial singularity (quite the reverse, in fact!), so the white hole type of initial singularity should be a common feature of unconstrained big bangs even if it does not result in an isolated white hole persisting with time.

The view is sometimes expressed, as an attempt to preserve time-symmetry, that the universe should finally return to a (fairly) low entropy state in the big crunch, the normal arrow of time being reversed in the collapsing phase of the universe (Gold 1962; Wheeler 1979). It may be pointed out that, in addition to all the other extreme difficulties that this viewpoint encounters (cf. Davies 1974; Penrose 1979) is the fact that if black holes are to be expected at all in such a universe model (and it is difficult indeed, to see how to avoid them) then a considerable population of white holes, and some very large ones too, would be expected to persist from the big bang until our present epoch and well beyond it – accompanied by highly correlated non-thermodynamic behaviour. I do not believe that the implications of such a model resemble, in the slightest degree, the universe as we know it, and I would challenge present supporters of this view to give a serious discussion of such matters!

Another issue which is often injected into the entropy–time-asymmetry discussion is the *anthropic principle*. The argument appears to date back to the time of Boltzmann (1895) and it asserts, in effect, that the low entropy that we observe about us is a consequence merely of our existence. The second law is, indeed, an essential ingredient of the workings of the metabolic processes upon which our lives depend, a high degree of order (i.e. low entropy) in our surroundings and in ourselves being necessary for our existence, in order that we may think, feel and perceive. Might it not be, so the argument runs, that only in those universes in which there is an operative second law and a high degree of order at some stage

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

(though perhaps only as part of a huge fluctuation), will the conditions suitable for life and consciousness arise?

The main difficulty with such a view is that only a very very tiny proportion of the 'specialness' of the initial state is actually of relevance to life. Indeed, since entropy is continually increasing, it would have been vastly 'cheaper', in terms of lowness of entropy, for the Creator merely to have selected a universe whose entropy was adequately low, say one million years ago, and to have allowed it to go up into the past before that time. In terms of the sorts of improbability that we have been discussing, the long tedious(?!) years of evolution which preceded this period were quite unnecessary! The fluctuation could simply have produced all the earthly living creatures intact! As a rough and extremely liberal calculation, for example, I estimate that all the particles in the solar system could simply 'happen' to come into their present positions with an improbability of one part in much less than $10^{10^{62}}$. If we require a healthy peppering of such inhabited solar systems throughout the universe – and let us be overwhelmingly generous by allowing every single star to have such a solar system, $\sim 10^{23}$ in all – then we obtain one in $10^{10^{85}}$ as an over-estimate of the improbability of it all happening simply by pure chance!

If, for example, we now allow the remaining inter-galactic gas to be collected into $10^{11} M_\odot$ black holes, then the Creator need only aim for a region of $\mathscr{W}$ whose volume (in absolute units) is $10^{10^{111}} \div 10^{10^{85}} \doteqdot 10^{10^{111}}$. Clearly this is a vastly 'cheaper' method than the one which appears actually to have been used (volume $\sim 10^{10^{88}}$)! And it is obvious, also, that one could improve considerably upon this figure. While I am happy to believe that most of $\mathscr{W}$ may well consist of uninhabited universes, because of a totally inhospitable chaotic geometry, there must nevertheless be a huge portion of it (with volume $10^{10^{115}}$ being surely a gross underestimate) in which conscious life abounds (assuming that physical laws not implying the hypothesis of Section 3 would allow life!) and in which the appearance of the universe is not at all close to the one in which we find ourselves, with its very closely Friedmannian beginnings (cf. Rees 1981).

I hope that it is clear from this that the anthropic principle cannot be what is responsible for the time-asymmetry of the second law. Indeed, it would appear from this that the Creator was not particularly 'concerned' about our existence, but was constrained in some very precise time-asymmetric way for some quite other reason. From this point of view, our present existence would arise merely as a by-product!

3. THE WEYL CURVATURE HYPOTHESIS

I have stated elsewhere (Penrose 1977; 1979) that a plausible hypothesis for the constraint on the big bang is that the Weyl conformal curvature vanishes as this singularity is approached. This would have the effect that 'gravitational clumping' is effectively absent in the initial state. The entropy in the matter can be assumed to be at a maximum, but the entropy in the gravitational field starts at zero. As the universe evolves, gravitational clumping increases, and this enables the entropy to rise above its original value, the greatest increases in entropy occurring when matter collapses into black holes.

The singularities of black holes, on the other hand, are not at all constrained in this way. In the spherically symmetrical case, the Weyl curvature diverges to infinity (according to the inverse cube of the radial distance) as the singularity is approached; and in the generic case, also, we expect the Weyl curvature to diverge to infinity at the singularity. In this respect, the 'final' singularities of black holes do not at all resemble the time-reverse of the structure that I am postulating for the big bang.

It would seem unnatural, however, merely to postulate one law for the big bang singularity and to have all other singularities unconstrained. It would seem that *all* 'initial' singularities (e.g. those occurring in white holes) ought to be constrained in the same way – and so I am claiming that they should have vanishing Weyl curvature – but this constraint should not apply to 'final' singularities. Indeed, if the *ideal point* definition of a singularity is used (Geroch, Kronheimer, and Penrose 1972; Penrose 1974) then singularities do in fact

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

appear as being of two types, namely the singular TIPs ('final' singularities) and singular TIFs ('initial' singularities). (See Fig. 3.1.) If strong cosmic censorship (Penrose 1974) is assumed then the TIPs and TIFs remain separated from one another, while without this hypothesis some singularities may have existence both as TIPs and TIFs.

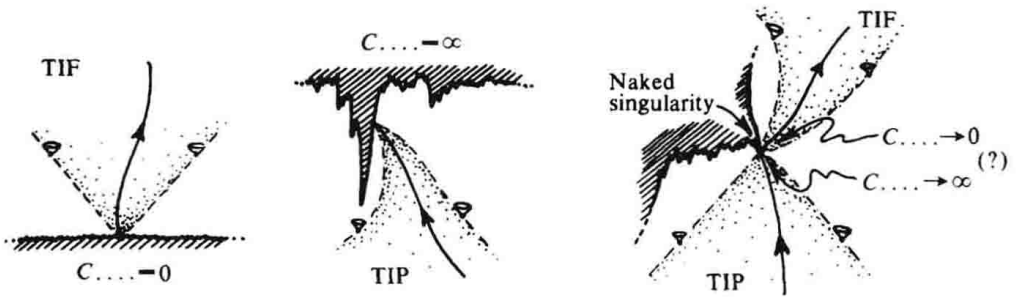


FIG. 3.1. A TIF describes an 'initial' singularity; and a TIP, a 'final' singularity. If strong cosmic censorship fails, a singularity can be represented both by a TIF and by a TIP.

Even when this occurs (such as in the presumed final ephemeral naked singularity of a Hawking-evaporating black hole) (Fig. 3.2) it seems not unreasonable that the curvature structure for the TIPs may differ from that for the TIFs (Penrose 1979). Accordingly, I am proposing:

Hypothesis: The Weyl curvature in any singular TIF tends to zero as the ideal point representing the TIF is directly approached from within the TIF.

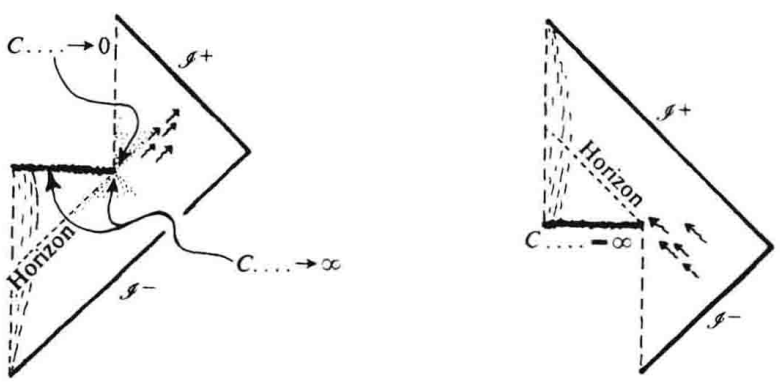


FIG. 3.2. Hawking-evaporating black hole (left) and its time-reverse (right).

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

Other slightly different statements could also be given (for example, the omission of the word 'singular' might be reasonable), and the statement is being kept deliberately a little vague. (The statement about approaching from within the TIP is being made in order that Hawking's ephemeral naked singularity should not necessarily violate the hypothesis.) See Figs. 3.1 and 3.2.

The manifold $\mathscr{W}$, that we considered earlier, must now be restricted to a tiny subset $\mathscr{U} \subset \mathscr{W}$. Thus, while in the definition of $\mathscr{W}$ we allowed all possible geometries for the initial singularity, the subset $\mathscr{U}$, to which the Creator's choices are now restricted, represents only those universes in which all initial singularities are constrained by the Weyl curvature hypothesis. Though the world-view that I am proposing (Penrose 1979) differs very greatly in its implications from that of chaotic cosmology, there is one essential feature of the chaotic cosmology viewpoint that I wish to retain (at least provisionally), namely that the Creator's pin is to be considered to be placed *at random* in the manifold of all allowed universe models – but in this case, in the manifold $\mathscr{U}$. This has the implication that the matter, in the early stages, is in thermal equilibrium, so one would anticipate that the Ricci tensor is initially very closely uniform, but that some (thermal) fluctuations are allowed – which of course, would be necessary in order that irregularities can arise. The exact nature and magnitude of these fluctuations would presumably ultimately depend upon some detailed particle physics.

It is an essential feature, also, that the final singularities should *not* be constrained in this way. The geometrical constraints on the initial geometry have the *effect* of producing a state of very low entropy – though I should emphasize that the Weyl curvature hypothesis is not in any way a statistical one, but is intended to be a precise physical law governing the structure of singularities (so it is not in any way necessary, for example, that a definition of gravitational entropy should be at hand). Likewise, any such constraints on final singularities would have the *effect* of producing a state of very low entropy in the future. That would have innumerable totally unacceptable physical implications, one of the clearest

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

being that all sorts of very precise correlations would exist in particle motions which would conspire to prevent any collapse to a black hole. These correlations could produce effects that would appear to us like pure magic, and physics would not at all resemble the physics that we know. The absence of such correlations in the initial state (e.g. the law of conditional independence of O. Penrose and Percival 1962) is, according to the world-view that I have been presenting, a consequence of the *absence* of constraints on final singularities, together with the randomness of the Creator's pin.

The other main implications of this world-view strongly involve the Weyl curvature hypothesis. These are:

(a) the second law of thermodynamics,*

(b) a closely Friedmannian early universe,†

(c) the absence of white holes,

and it would seem, also, that the absence of chaos in the early stages should preclude any appearance of black mini-holes.

It is my impression that one reason people may have felt uncomfortable with a big bang as uniform as the one that I am proposing, is that it *seems* to run counter to the appealing idea that all matter may have originated out of the vacuum, by means of particle-creation processes incurred by the initial highly curved state. If the Weyl curvature was initially zero (and if the big bang was indeed a curvature singularity), then all the initial curvature must have already resided in the Ricci tensor, i.e. by Einstein's equations, matter must always have been present right from the start! However, there is no reason (provided that we can accept non-conservation of baryon number) that the initial 'matter' should not have been something with the quantum numbers of the vacuum. The initial Ricci tensor *could*, perhaps, have been entirely an 'effective energy tensor', produced because vacuum fluctuations give rise to an effective modification of Einstein's equations in regions of very high curvature.

---

* The retardation of radiation is one aspect of this.

† Exactly *how* closely Friedmannian it is, would depend upon the statistical uniformity of the Ricci tensor.

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

Another reason that people may find the Weyl curvature hypothesis hard to accept is that it is grossly time-asymmetric. Apart from K^0-decay, there seems to be nothing else in the local laws of physics which violates time-symmetry. However I would insist that to explain the second law of thermodynamics we do need a grossly time-asymmetric law of physics (cf. Section 2). The local laws of physics that we know and understand are, it is true, all time-asymmetric. But the very presence of space-time singularities in realistic models, together with the fact that these singularities are 'regions where the presently known laws of physics break down', suggests that we are in need of laws *other* than those that we know and understand. These new laws are required in order that the physics of what we now refer to (in the absence of a proper theory) as 'space-time singularities' may eventually be understood.

4. HAWKING'S BLACK = WHITE HYPOTHESIS

Let me now turn to another quite different point of view which has been put forward by Hawking (1976*a*,*b*). Again, the starting-point is the Bekenstein–Hawking formula, but employed in a quite different context. Hawking envisages a box of volume V, with perfectly reflecting walls, in which there is a certain amount, E, of energy in some form. The box is left undisturbed for an indefinite period and its contents allowed to find equilibrium. Now provided that E is large enough (but not so large as to collapse the entire box), given V, we find that the maximum entropy is achieved when the box contains a black hole in equilibrium with a small amount of thermal radiation. According to Hawking's argument, the physics which has entered into the discussion of this problem is entirely time-symmetric, namely general relativity, the standard formalism of quantum field theory, and Maxwell theory. Thus, it is claimed, the configuration of maximum entropy ought also to be time-symmetric. (Hawking also presents another argument, based on the second law. A modified version of this will have its role to play in the next section.) Since the maximum entropy configuration is a black hole, with radiation, and

since the time-reverse of a black hole is a white hole, then the conclusion is that black holes and white holes are physically indistinguishable!

Alarming as such a suggestion may seem, at first, there is a certain elegance and economy in the idea. The Hawking radiation emitted from a black hole can now be re-interpreted as matter-creation at the white hole's singularity. Likewise, the matter or radiation which the black hole swallows may be re-interpreted as the time-reversed Hawking radiation from the white hole. The apparent difference between the random Hawking radiation and the coherent objects (e.g. television sets) that can be swallowed by black holes is explained (reasonably convincingly) in terms of the time-asymmetric boundary conditions that are normally imposed on external objects (that is, it is normally considered reasonable to impose a low-entropy input, e.g. a television set, but not to select for a low-entropy output).

However, I have never found myself able to accept this view that black holes and white holes should be physically the same. A glance at Fig. 3.2 should make clear what are my difficulties in this respect. The left-hand diagram represents the formation of a black hole by classical gravitational collapse followed, ultimately, by its disappearance as Hawking radiation; the right-hand diagram represents the formation of a white hole by time-reversed Hawking radiation, followed by its disappearance by the classical disgorging of some ordinary matter. Even if the entire portions inside the horizons are deleted, we obtain two space-time geometries which are quite different. Hawking's answer to this criticism is that it is not appropriate to insist upon the existence of a uniquely defined objectively determined space-time geometry when considering processes of this nature. Since the back-reaction on the space-time, of the quantum mechanical Hawking radiation, is an essential ingredient of these pictures, the space-time structure has itself become subject to quantum-mechanical rules. Thus, Hawking envisages that the space-time geometry does not have objective existence but has become an essentially observer-dependent construct.

For myself, I find this line of argument very hard to accept. It seems to me that if we are talking about reasonably large

black holes of a solar mass or more, then the space-time geometry ought to be sufficiently well-defined that an objective general-relativistic picture is valid and that the black hole and white hole ought to be distinguishable by physical experiment. As a further objection, it had seemed to me clear that the most probable mode of formation of a black hole in the box (by classical collapse) was not at all the time-reverse of its most probable mode of disappearance — even if all that one was allowed to 'observe' was the radiation reaching (or leaving) the walls of the box. However, these arguments have become involved in much dispute, the main difficulty being that no-one seems yet to have been able to perform the relevant calculations of the probabilities involved in a sufficiently trustworthy manner.

A problem is that one is dealing with relaxation times that are ridiculously long when the mass is, say $M_\odot$ or more, if we wish to wait for the hole to disappear and to reform merely by chance fluctuations. Even if $V = \infty$, it would take 10^{56} Hubble times, or so, for a $M_\odot$ black hole to disappear, and in this case it will almost certainly never reform. For a smaller box containing $M_\odot$, with V say the size of a galaxy, the black hole would represent maximum entropy, so its occasional disappearance by a fluctuation would take far far longer. Also, for it to form again out of radiation would involve time-scales and a precise mode for formation that are not easy to estimate in any detail. I do not have reason to doubt the qualitative arguments against time-symmetry that I have presented earlier (Penrose 1979), but I accept that there is considerable room for argument and that the discussion is perhaps not so conclusive as I had previously thought.

In Fig. 4.1, I have indicated the three different situations: (i) large box with pure radiation as maximum entropy and unstable hole, (ii) medium-sized box with a locally stable hole but pure radiation still as maximum entropy and (iii) smallish box with the hole as maximum entropy (cf. Hawking 1976*a*,*b*; Gibbons and Perry 1978). In Table 4.1 are listed various familiar objects, providing the mass-energy E and the corresponding volumes V for which the maximum entropy is about equally balanced between the pure radiation state and the

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

black hole, so that some feeling may be obtained for the scales involved.

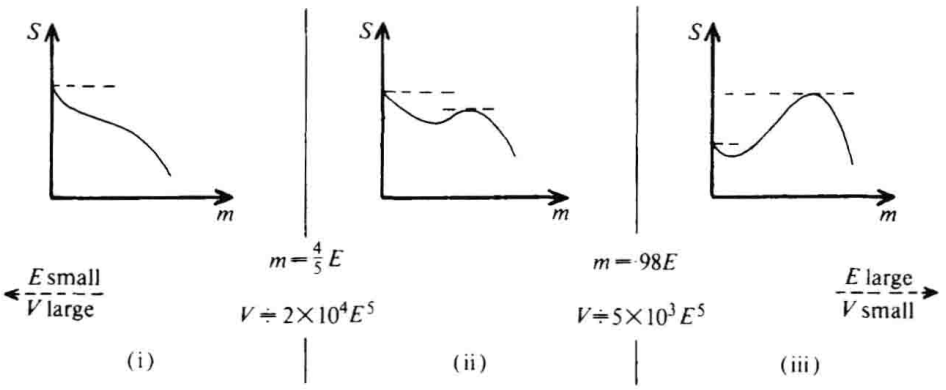


FIG. 4.1 i-iii. The entropy S plotted against the mass m which is contained in a black hole, for total energy E in a box of volume V ($c = \hbar = G = k = 1$).

TABLE 4.1

Volumes V and mass-energies E for which pure radiation or a black hole have roughly equal entropies

(Note: a black hole of 10^{15}g has roughly the radius of a proton; but the mass of a small mountain. $M_\oplus$ is the mass of the Earth.)

E	V
pyramid	10^{-7}cc
10^{15}g	10^{3}cc (ordinary box)
comet	Vol. of Moon
$M_\oplus \doteq 6 \times 10^{27}$g	10^{67}cc (galactic dim.)
$M_\odot \doteq 2 \times 10^{33}$g	10^{95}cc (bigger than obs. universe)

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

5. SELF-COLLAPSING WAVE FUNCTIONS?

From my own point of view, Hawking's suggestion that black and white holes might be physically the same would not be acceptable for another reason. For the Weyl curvature hypothesis of Section 3 rules out white holes as having unacceptable singularity structure, whereas it allows black holes. In fact, if this hypothesis is admitted, it invalidates the argument that the physical laws involved in the discussion of Section 4 are all time-symmetric. Part of the physics involved inside the box is the physics of space-time singularities, since this is a necessary part of the discussion of black or white holes. Thus, I would maintain that when maximum entropy is achieved by the black hole state, this state need not (and, indeed *cannot*) be time-symmetric, since some time-asymmetric physics (namely the Weyl curvature hypothesis) is crucially involved in the discussion.

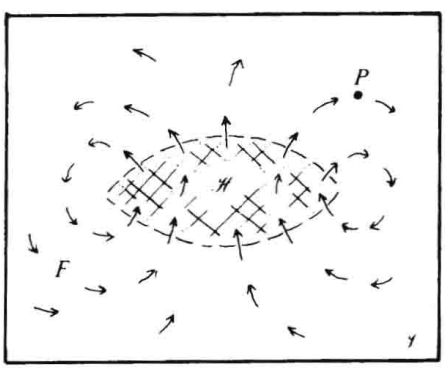
However, this does not absolve me from considering Hawking's other objection to time-asymmetry (Hawking 1976*a*,*b*) which was based on considerations of the second law of thermodynamics. I prefer, however, not to consider Hawking's argument in its original form (since the second law's validity seems to me to be somewhat questionable in any case, when time-scales as long as the ones needed here are involved). Instead I shall give a related argument which seems to me to contain the essence of Hawking's objections.

Consider the phase space $\mathcal{V}$ of the contents of the box. I shall be somewhat vague as to whether this is to be a classical phase space or a quantum one. If classical, it is a space of finite volume. If quantum-mechanical, it is a finite-dimensional Hilbert space. This is because the volume V and energy E are both bounded finite quantities, the standard rules of classical or quantum mechanics being assumed to apply. Let us now consider the flow F on the space $\mathcal{V}$, which gives the time-evolution of the contents of the box. It will be assumed that it is valid to consider an essentially non-relativistic description, with the contents of the box given as a function of time – even when the box contains one or more black or white holes. As the system evolves, the corresponding phase-space point P moves

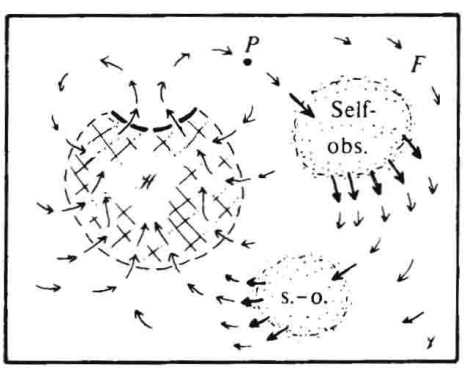
Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

about $\mathcal{V}$, as determined by the flow F. At certain times the box will contain at least one hole, either black or white (or both), and at those times the point P lies in a special subset of $\mathcal{V}$ labelled $\mathcal{H}$. (I am ignoring any difficulty about defining the precise moments at which the box is considered to contain such a hole, since this does not seem to have much effect on the discussion.)

Now there are various viewpoints that one may consider. One of these is Hawking's, according to which black and white holes are physically the same and there is an essential time-symmetry in the whole set-up. Another, in which there is also time-symmetry, is the 'common-sense' viewpoint, according to which black holes can exist and white holes can exist but they are simply different objects from each other, and furthermore many 'zebroid' combined holes might also exist. In each of these viewpoints, because of the time-symmetry, there are 'as many' flow lines entering the region $\mathcal{H}$ as there are leaving it. The situation is depicted in the left-hand diagram of Fig. 5.1.



Hawking viewpoint, or 'common-sense' viewpoint



A viewpoint involving the Weyl curvature hypothesis

FIG. 5.1. The phase space $\mathcal{V}$ of the contents of the box, $\mathcal{H}$ being the region of $\mathcal{V}$ for which the box contains at least one hole (black or white or both).

My own viewpoint is depicted in the right-hand diagram, and here there are 'more' flow lines entering $\mathcal{H}$ than there are leaving it. The reason for this is that there are, in my viewpoint, many more ways of creating a black hole than there are ways of destroying it. For a black hole can only disappear

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

quantum-mechanically by means of the Hawking process, while it can be produced either quantum mechanically or by means of classical collapse. Since a black hole is not a white hole, it cannot simply disappear by means of those evolutions which represent the time-reverse of collapse. Those motions are not physically allowed since they involve a space-time singularity which does not satisfy the Weyl curvature hypothesis.

Another way of roughly seeing this is to look again at Fig. 3.2. Only the left-hand diagram is allowed, while to specify data for geometries of that type, information may be given* on $\mathscr{I}^-$, or equivalently *both* on $\mathscr{I}^+$ *and* on the singularity. Thus, while the data on $\mathscr{I}^-$ may be freely specified, only a restricted class of data on $\mathscr{I}^+$ is allowed because this must be supplemented by extra data on the singularity in order to give as many states as those allowed on $\mathscr{I}^-$. (Indeed, it was considerations of this kind which led Hawking to discover his radiation in the first place.) The disallowed data on $\mathscr{I}^+$ would, when evolved backwards, give geometries resembling that of the right-hand diagram of Fig. 3.2, or also with final singularity data, geometries resembling combinations of both diagrams in Fig. 3.2. All of these are excluded in my viewpoint.

It is here that my viewpoint runs into some trouble. According to Liouville's theorem the phase-space volume should be preserved as we follow the flow. However, phase-space volume is lost within $\mathcal{H}$ because 'fewer' flow lines leave $\mathcal{H}$ than enter it (more accurately, there is a net loss of volume per unit time leaving the region $\mathcal{H}$). There is no difficulty about contradicting Liouville's theorem for the region $\mathcal{H}$ since for that region a space-time singularity exists within the box so we expect that the ordinary laws of physics must be partially suspended. The essential difficulty comes with the region *outside* $\mathcal{H}$. In *that* region normal physics is supposed to hold good, so it would come as more of a shock if we were to be

---

* Strictly speaking, there is also some information in the final singular TIF of the Hawking-evaporated black hole, but (at least in my own viewpoint) this information counts for very little and will be ignored in this discussion.

confronted with a contradiction to Liouville's theorem there. But if we *assume* that Liouville's theorem holds good in $\mathscr{V} - \mathcal{H}$ we seem driven to conclude that certain regions within $\mathscr{V} - \mathcal{H}$ must become inaccessible once $\mathcal{H}$ has been cycled through. These regions, in fact, correspond to physical motions that, when extrapolated backwards, would lead to time-reversed gravitational collapse in which there is a singularity violating the Weyl curvature hypothesis (e.g. to white holes). These regions need not be excluded when the box and its contents are *first* formed, but seem to be disallowed as part of whatever 'equilibrium' configurations the contents of the box finally settle into. This gives us a rather strange picture of the ultimate thermodynamics of the contents of the box.

One can also look at this difficulty from the quantum-mechanical point of view. Here Liouville's theorem translates into conservation of probability. If no 'observations' take place, the Schrödinger evolution holds good and single pure states are supposed to evolve into single pure states. The generalization of this view given by Hawking (1976*b*) to cover his own viewpoint allows single pure states to evolve into density matrices. This, in effect, allows flow lines to 'spread' into wider regions, but a probability weighting is given, so that the total spread-out region counts just the same as a single line. In my own viewpoint, then, there is classically a volume-reduction taking place whenever $\mathcal{H}$ is cycled through, and this is interpreted quantum-mechanically as a reduction in the (finite) dimension of the Hilbert space of allowable states, as compared with the dimension of the space of states initially available. In neither Hawking's viewpoint nor the 'common sense' one does the difficulty that I am encountering with my own seem to occur. My objections to those other viewpoints do not spring from the present considerations but from all the other several difficulties that allowing white holes to exist seem to involve us in.

However, some of those other difficulties do, in my opinion, provide us with a clue. In Hawking's viewpoint, the geometry begins to get inextricably involved with quantum mechanics as the system evolves. (Recall the 'observer dependence' of the

geometry that we seemed to be forced into.) After the box has been left in isolation for a very long time, its contents would not resemble a classical geometry at all. It seems that an 'observer' would be needed to peer into the box and so collapse the state into that of a well-defined geometry. I have always had difficulties with that aspect of quantum mechanics. It seems to me that something objective must actually be taking place within the box when such gross features such as the existence or location of a huge black hole are being considered. Perhaps quantum linear superposition fails when space-time geometry is involved? Indeed, that is the kind of line that I am proposing to take. My point of view is not so unconventional as it may seem at first. The normal Copenhagen interpretation requires, after all, that 'Rule 1' (Schrödinger evolution of the state vector) is sometimes replaced by 'Rule 2' (discarding the old state vector and instituting a new one which is an eigenstate of an appropriate operator, these being the times when the quantum mechanical probabilities arise). It is nowhere stated exactly *when* Rule 2 is to be used, but it involves the presence of some large 'classical' structure. Might it not be (cf. Karolyhazy 1966; Komar 1969) that this rule comes into play when systems become sufficiently 'macroscopic' that they significantly affect the space-time geometry?

There is a relation also to entropy. An 'irreversible process' is normally considered necessary in order that Rule 2 be brought into play. This would be stated as: 'the entropy of the system has gone up'. But entropy is normally considered to be a rather subjective concept, dependent upon someone's notion of coarse-graining. It is hard to see how an objective change in the state of the world can take place according to Rule 2 unless 'entropy increase' can sometimes be objective. The best example in physics, so far, in which an entropy seems to have acquired an objective meaning, is in the Bekenstein–Hawking formula. Here the entropy is equated to a precise space-time-geometric quantity, namely an absolute event horizon's surface area. So if space-time geometry is to be regarded as in some sense objective, then entropy also can apparently acquire objectivity in these circumstances.

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

I regard these considerations as representing encouragement for the following picture: for much of the time the evolution of a system can be very accurately treated as proceeding according to Rule 1. But when macroscopic space-time geometry becomes inextricably involved, Rule 2 comes objectively into play as an accurate representation of reality. The precise boundary between the application of these two rules appears to be effectively unimportant (von Neumann 1955) – though ultimately we shall require some new theory which tells us what 'really' happens. No 'conscious observer' is required to intervene. Rule 2 just takes place *spontaneously* when the circumstances of the coupling of the system to space-time geometry are appropriate. I regard this picture as an approximation to some undiscovered theory of 'quantum gravity' in which both the Rules 1 and 2 arise as approximations to this more correct, but yet undiscovered theory. Thus, my own view of what the term 'quantum gravity' means appears to differ from that of most physicists. In particular, I believe that the correct quantum gravity must be a *time-asymmetric theory!*

Let us return to the discussion of the right-hand diagram of Fig. 5.1. What is required, in order that we can obtain a reasonable phase-space picture, is that exactly compensating for the loss of phase-space volume in $\mathcal{H}$ must be other regions of $\mathscr{V}$ in which phase-space volume is regained. These regions I envisage as representing the transitions from the application of Rule 1 to Rule 2 and then back to Rule 1 again. So long as Rule 1 applies, we follow along a unique trajectory of the flow F. But then, somewhere along the line, *self-observation* of the system in the box takes place, so that Rule 2 operates. A number of different trajectories emerge from this rather vague region in $\mathscr{V}$, these representing the different possible outcomes of the result of this self-observation. The system could, indeed, 'select' any one of these outcomes. This is where the non-determinism of quantum mechanics comes in. The system has to 'make its choice' and thereafter it continues following some definite trajectory of the flow F (Rule 1 back in operation again). Now, after wandering round and round the space $\mathscr{V}$ for an enormous length

of time, our phase-space point P will finally come back in along the same trajectory that we considered before. The self-observing system is confronted with exactly the same 'choice' as before, but now it may well be that a different selection is made, and P goes wandering off along some quite different path. Each time, the point comes in along some definite path and it goes out along some definite path. The (quantum mechanical) Liouville theorem has the *appearance* of not being violated. Yet because many more alternatives leave the 'self-observation' regions than enter them we do have an *effective* net volume increase, in violation of Liouville's theorem.

I shall try to be a little more explicit about what is involved, in this picture, when Rule 2 comes into play. I have asserted that whenever the region $\mathcal{H}$ is cycled through, the dimension of the Hilbert space of allowable states gets reduced, the non-allowable states being those which when evolved backwards would lead to geometries involving a violation of the Weyl curvature hypothesis. Now, as time progresses, it will become harder and harder to ascertain precisely which these non-allowed states are: an initial specialness of geometry (an apparently low entropy) will develop into correlations between the detailed particle motions (an apparently higher state of entropy). The Hilbert space $\mathscr{V}$ of allowable states may be thought of as a subspace of a larger space $\mathscr{Y}$ for which the Weyl curvature constraint is not imposed. (Compare the corresponding spaces $\mathscr{U} \subset \mathscr{W}$ of Section 3.) We may regard $\mathscr{Y}$ as a Hilbert space of constant dimensionality, while $\mathscr{V}$ now has a dimension that changes with time. After a Hawking evaporation, the subspace $\mathscr{V}$ is partly discernible in terms of macroscopic features of the geometry*, but this becomes progressively less so with time, as some gross macroscopic features of geometry get 'lost' in particle correlations. At some stage Rule 2 comes into operation as a consequence (?) of this and the

* The point of final explosion is clearly a state of less than maximum entropy. This is not unreasonable since a large fluctuation was needed at an earlier stage in order for this situation to arise.

Time-asymmetry and quantum gravity. *Quantum Gravity 2: A Second Oxford Symposium.* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 244–72. 1981

state vector, which initially lies in the subspace $\mathscr{V}$ (because of the Weyl curvature hypothesis) gets projected into some state vector in $\mathscr{Y}$ with components lying outside $\mathscr{V}$. (The 'natural' or 'macroscopically discernible' Hilbert space bases in $\mathscr{Y}$ are no longer specially related to $\mathscr{V}$.) The space of allowable states has now grown in dimensionality – and we have to *relabel* this larger subspace of $\mathscr{Y}$ as our new $\mathscr{V}$. This dimensionality increase is just the net volume increase that we require, in an effective violation of Liouville's theorem.

The idea, of course, is that this volume increase should exactly compensate for the volume loss in the black hole region $\mathcal{H}$. In fact we now have the intriguing possibility of actually performing a calculation which estimates the amount of self-observing that a system of energy E in a volume V ought to be indulging in, according to this point of view. Since the formation of black holes is an entirely gravitational phenomenon, we now have, in principle, a quantitative link between gravity, as represented by black holes, via the Bekenstein–Hawking formula, and the quantum-mechanical observation process. It seems that all one need do is to calculate the volume loss in the region $\mathcal{H}$, for given E and V, and use this to obtain a measure of averaged effective volume-spreading due to self-observation. An appealing aspect of this is that while it is obvious that black holes need not be present in any particular observation process, their theoretical existence *is* necessary, in this picture of things, to produce the correct overall balance.

All this will repay closer examination. But as things stand, I have found considerable difficulty in estimating what the suggested self-observation rate ought to be. The calculation of the volume-loss within the region $\mathcal{H}$ is complicated by the fact that there is also some volume-gain within $\mathcal{H}$ owing to applications of Rule 2 whilst black holes are present. There is also something a little disturbing, in this respect, about Table 4.1. The mass-energy contained in a pyramid is certainly ample for a laboratory and a team of physicists performing quantum interference experiments. If they were suddenly encased in a box with perfectly reflecting walls, they would certainly be able to continue to perform such experiments for a long

while to come, presumably employing many applications of Rule 2 before finally expiring. Is it reasonable to suppose that this needs to be compensated for by the occasional collapse of the mass-energy in their (sometime) laboratory into a black hole considerably smaller than a proton? This is very very far from the maximum entropy configuration since here $V \gg 10^{-7}$cc.

This all sounds rather absurd, yet I feel that it does not invalidate the earlier discussion. The observations which are taking place when the box is first formed may have more to do with the big-bang specialness $\mathscr{U} \subset \mathscr{W}$, than $\mathscr{V} \subset \mathscr{Y}$. But perhaps one must also be much more careful about the idealizations involved in the idea of a box with perfectly reflecting walls. In practice – and also in principle – there are severe difficulties with that concept. Our physicists in the above discussion could *not* in fact be isolated, by such a box, to the required degree.

These arguments remain inconclusive as they stand. I do, however, strongly believe that any successful quantum gravity theory will have something to say about such matters. Space-time singularities, thermodynamical time-asymmetry and quantum mechanical observations all seem to be interlinked with quantum gravity, the Bekenstein–Hawking formula forming one cornerstone of present discussion of such matters.

ACKNOWLEDGEMENTS

I am grateful to Stephen Hawking, Bob Wald and a number of other colleagues for many inspiring conversations. Most particularly, I am indebted to Don Page for some penetrating criticisms and also for supplying many of the numerical values that I have used.

REFERENCES

Barrow, J.D. and Matzner, R.A. (1977). *Mon. Not. R. astr. Soc.* **181**, 719.
Bekenstein, J.D. (1973). *Phys. Rev.* **D7**, 2333.
—— (1974). *Phys. Rev.* **D9**, 3292.
Boltzmann, L. (1895). *Nature*, **51**, 413.
Davies, P.C.W. (1974). *The Physics of Time Asymmetry*. Surrey University Press.

Eardley, D.M. (1974). *Phys. Rev. Lett.* **33**, 442.
Geroch, R., Kronheimer, E.H., and Penrose, R. (1972). *Proc. R. Soc.* **A327**, 545.
Gibbons, G.W. and Perry, M.J. (1978). *Proc. R. Soc.* **A358**, 467.
Gold, T. (1962). *Am. J. Phys.* **30**, 403.
Hawking, S.W. (1975*a*). *Comm. Math. Phys.* **43**, 199.
—— (1975*b*). In *Quantum Gravity, an Oxford Symposium.* (ed C.J. Isham, R. Penrose, and D.W. Sciama). Oxford University Press.
—— (1976*a*). *Phys. Rev.* **D13**, 191.
—— (1976*b*). *Phys. Rev.* **D14**, 2460.
Karolyhazy, F. (1966). *Nuovo Cim.* **A42**, 390.
Komar, A. (1969). *Int. J. Theor. Phys.* **2**, 157.
Littlewood, J.E. (1953). *A Mathematician's Miscellany.* Methuen, London.
Misner, C.W. (1968). *Astrophys. J.* **151**, 431.
—— (1969). *Phys. Rev. Lett.* **22**, 1071.
Neumann, J. von (1955). *Mathematical Foundations of Quantum Mechanics.* Cambridge University Press.
Penrose, O. and Percival, I.C. (1962). *Proc. R. Soc.* **79**, 509.
Penrose, R. (1974). In *Confrontation of Cosmological Theories with Observational Data (IAU Symp. 63).* (ed M.S. Longair). Reidel, Boston.
—— (1977). In *Proc. First Marcel Grossmann Meet. Gen. Relativ. (ICTP Trieste).* (ed R. Ruffini). North-Holland, Amsterdam.
—— (1979). In *General Relativity, An Einstein Centenary Survey.* (ed S.W. Hawking and W. Israel). Cambridge University Press.
Ramaswamy, S. and Wald, R.M. (1980). *Phys. Rev.* **D21**, 2736.
Rees, M.J. (1981). (This volume).
Wheeler, J.A. (1979). *Frontiers of Time* (North Holland, for Società Italiana di Fisica, Amsterdam).
Zel'dovish, Ya.B. (1972). In *Magic Without Magic: John Archibald Wheeler.* (ed J. Klauder). Freeman, San Francisco.
—— (1974). In *Gravitational Radiation and Gravitational Collapse (IAU Symp. 64).* (ed C.M. De Witt). Reidel, Boston.

122
SOME REMARKS ON TWISTORS AND CURVED-SPACE QUANTIZATION

Quantum gravity 2: A second Oxford symposium.
Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92.
Clarendon Press, Oxford, 1981

Twistor ideas are used in an approach to the formulation of quantum field theory in general curved space-times, using ideas due to Andreotti and Hill, and others, concerning sheaf cohomology of CR hypersurfaces. Fundamental difficulties remain unresolved, however—not unexpectedly. (I have a nagging feeling that an acknowledgement to George Sparling is due, in this paper. I believe that he had independent ideas similar to those expressed here, but my memory is too unclear for me to be specific.)

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

SOME REMARKS ON TWISTORS AND CURVED-SPACE QUANTIZATION

R. Penrose
Mathematical Institute, Oxford University

1. INTRODUCTORY REMARKS

Let me start by making some general comments concerning twistor theory (Penrose 1967, 1975; Penrose and MacCallum 1972; Penrose and Ward 1980; Hughston and Ward 1979). Some people at this conference might perhaps be puzzled that the theory seems to be trying to play two diametrically opposite roles. An impression that may have been gained from Chris Isham's clear and comprehensive opening survey, for example (Isham in this volume), is that twistor theory is something possibly to be resorted to if all else fails: if our accepted physical pictures of space-time and of quantization procedures finally land us with irresolvable contradictions, then perhaps (before giving up altogether!) we might (briefly?) turn to twistor theory, as we seem driven to seek radical alternative viewpoints as regards the ultimate nature of space and time. On the other hand, in Richard Ward's elegant lecture (Ward in this volume), twistor theory was seen in quite a different light, namely as something that has a mathematical role to play as an aid in the solution of certain systems of partial differential equations of some possible interest to physics, twistors being, in effect, regarded as a source of 'mathematical tricks', but without necessarily having any deep significance in terms of physics.

In attempting to explain this apparent conflict of purpose, I should begin by stressing that, while twistor theory is undoubtedly striving to do something 'new' as regards one's descriptions of space-time, it does *not* throw overboard existing physical theory. Twistor concepts are firmly based in established physics – notably those parts of physics (special relativity, Maxwell theory and the basic formalism of quantum mechanics) that have become most forcefully established in

experiments performed over the past century. Twistors present, at least in the first instance, merely a mathematical reformulation of physics, so, in a clear sense, the excellent experimental support that this physics enjoys is support also for twistor theory! What is new in the twistor approach (and when I say 'new', I mean merely 16 years old) is a shift in emphasis in what are to be regarded as the primary elements of physical theory. Thus, while in the normal view space-time points are the primary ingredients in terms of which things are to be described, according to the twistor view, space-time points are *derived* rather than primary objects. So, whereas the twistor picture is initially equivalent to the normal space-time view (at least as regards special relativity), it offers scope for generalizations in which this ceases to be the case, where space-time points have become 'fuzzy' objects or have perhaps disappeared altogether.

The dual nature of the role of twistor theory should be clear from this. Inasmuch as it is equivalent to conventional physical theory, twistor theory provides us with a mathematical reformulation of the equations of physics – at least those which have so far been successfully transcribed – often yielding novel insights that are hard to come by in a more direct approach. In so far as the theory suggests new developments for physics, it does indeed offer scope for a radical change, the nature of which would differ in principle from any change that could plausibly be anticipated from the standpoint of conventional space-time theory. It should be emphasized, however, that while such a change is very much a part of the motivation behind the twistor approach, almost all of the work that has so far actually been carried out in the subject has been concerned with mathematical reformulations. The idea is that once an entirely twistorial description of some part of physics has been achieved, then the natural road to its most promising generalizations should become apparent!

I ought to make one comment by way of qualification of some of the above remarks. This concerns the twistor transcription of *general* relativity. It may be said that twistor theory owes much to Einstein's view of gravity. The idea of covariance, of the geometrization of physics, and of the importance of

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

allowing the basic manifolds of the theory to become curved; all these have had essential roles to play in twistor theory. And in more than one place in the development of the theory, key ideas have arisen out of certain partially successful attempts to transcribe general relativity. Yet Einstein's theory has so far remained stubbornly resistant to a complete transcription into twistor terms. At present, it is not clear that a satisfactory transcription is possible at all, although in my view (not shared by all my colleagues!) there are some remarkably promising indications. My earlier comments about physical laws being in principle re-expressible in twistor terms do not, strictly speaking, apply to general relativity as things stand.

2. STANDARD TWISTOR DESCRIPTIONS

Let us recall, first, the standard flat-space twistor correspondence. The geometry is clearest in terms of the *projective* twistor space $\mathbb{PT}$, whose points are labelled by homogeneous complex coordinates $Z^0:Z^1:Z^2:Z^3$, where the Zs are the components of a twistor Z^α. A point of Minkowski space, with standard Minkowski coordinates t,x,y,z (metric diag (+1,-1, -1,-1)) is said to be *incident* with the twistor Z^α whenever the equation

$$\begin{pmatrix} Z^0 \\ Z^1 \end{pmatrix} = \frac{i}{\sqrt{2}} \begin{pmatrix} t+z & x+iy \\ x-iy & t-z \end{pmatrix} \begin{pmatrix} Z^2 \\ Z^3 \end{pmatrix} \tag{2.1}$$

holds. This equation may also be written in the 2-spinor form

$$\omega^A = ix^{AB'}\pi_{B'}, \tag{2.2}$$

where $(\omega^0,\omega^1,\pi_{0'},\pi_{1'}) = (Z^0,Z^1,Z^2,Z^3)$ and where $(x^{00'},x^{01'},x^{10'},x^{11'}) = 2^{-\frac{1}{2}}(t+z,\ x+iy,\ x-iy,\ t-z)$. The condition for (2.1) or (2.2) to have solutions in real Minkowski space $\mathbb{M}^I$ (i.e. t,x,y,z real, i.e. $x^{AB'}$ Hermitian) for a given twistor Z^α (with $\pi_{A'} \neq 0$), is

$$Z^\alpha \bar{Z}_\alpha = 0$$

where $(\overline{Z}_0, \overline{Z}_1, \overline{Z}_2, \overline{Z}_3) = (\overline{Z^2}, \overline{Z^3}, \overline{Z^0}, \overline{Z^1})$. The twistor norm $Z^\alpha \overline{Z}_\alpha$ can also be written

$$Z^\alpha \overline{Z}_\alpha = Z^0\overline{Z^2} + Z^1\overline{Z^3} + Z^2\overline{Z^0} + Z^3\overline{Z^2}$$
$$= \omega^A \bar{\pi}_A + \pi_{A'}\bar{\omega}^{A'}$$

which is a Hermitian form of signature (++--). The solutions for $x^{AB'}$ in M^I of (2.1) or (2.2) constitute a null straight line whenever $Z^\alpha \overline{Z}_\alpha = 0$, $\pi_{A'} \neq 0$. If $\pi_{A'} = 0$, we also say that Z^α determines a null straight line, but this time a generator of $\mathscr{I}$, the light cone at infinity which completes $\mathbb{M}^I$ to a compact manifold $\mathbb{M}$. The real 5-dimensional hypersurface in $\mathbb{PT}$ that is given when $Z^\alpha \overline{Z}_\alpha = 0$ is called $\mathbb{PN}$.

We have just seen that the points of $\mathbb{PN}$ represent null geodesics in $\mathbb{M}$. In the reverse correspondence, any point of $\mathbb{M}$ represents a projective line in $\mathbb{PT}$ which lies entirely within $\mathbb{PN}$. This may be seen from the incidence relations (2.1) and (2.2) by holding the Minkowski point fixed and allowing Z^α to vary. The space of null geodesics through a fixed Minkowski point (i.e. the generators of its light cone) correspond in $\mathbb{PN}$ to this projective line (Fig. 2.1). The topology of each

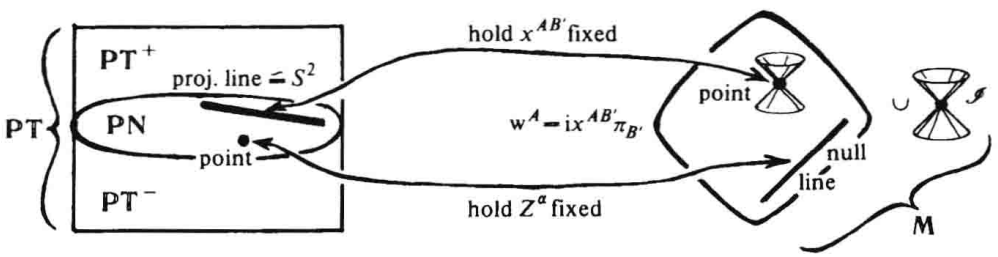


FIG. 2.1. The standard correspondence between $\mathbb{PN}$ and $\mathbb{M}$

such a projective line is a sphere S^2.

We also use the correspondence (2.2) or (2.1) in the case of *complexified* points ($x^{AB'}$ not necessarily Hermitian). Holding $x^{AB'}$ fixed we still get a projective line in $\mathbb{PT}$ to represent it, but where the line does not necessarily lie in $\mathbb{PN}$. In fact it lies entirely in $\mathbb{PT}^+$ ($Z^\alpha \overline{Z}_\alpha > 0$) or $\mathbb{PT}^-$ ($Z^\alpha \overline{Z}_\alpha < 0$) if the complexified Minkowski point lies in the *forward* or *backward tube*, respectively. (Holding Z^α fixed, we now find that the complexified Minkowski points which satisfy (2.1)

or (2.2) constitute a type of totally null plane referred to as an α-plane.)

Note that the correspondence is a non-local one. A small, say convex, region in (complexified) space-time is represented by a non-local region in $\mathbb{PT}$ (swept out by projective lines) whose topology has the homotopy type of a sphere S^2. If we were to choose a 'smaller' region of $\mathbb{PT}$ not containing such a suitable topological sphere, or else some region of $\mathbb{PT}$ that is, in some sense, too 'bent' (so that none of the spheres that it contains has the holomorphic structure of a projective line) then the space-time 'points' that it attempts to represent would have somehow 'disappeared'. If we were to deform the twistor space itself, we might find that it then contains *no* global region which is sufficiently extensive to enable us to define a space-time point – even though the *local* holomorphic structure of the (projective) twistor space may be identical with that of $\mathbb{PT}$. Such 'deformed' twistor spaces in fact, play an essential role in the twistor description of general relativity (Penrose 1976), but where, in the normal situation, the deformations are not so great that the corresponding space-time points cease to exist.

3. MASSLESS FIELDS AND SHEAF COHOMOLOGY

The standard correspondence, whereby (massless) fields in space-time are defined in terms of holomorphic 'twistor functions', is also an essentially non-local one. We envisage that the relevant portion of twistor space (say $\mathbb{PT}^+$, for positive frequency fields) is covered by a locally finite system of open sets $\mathcal{U}_i$, the 'twistor function' being a system of holomorphic functions f_{ij} defined on the intersections $\mathcal{U}_i \cap \mathcal{U}_j$ of *pairs* of these open sets. The fs satisfy

$$f_{ij} = -f_{ji} \tag{3.1}$$

and are subject to the *cocycle condition*

$$f_{ij} + f_{jk} + f_{ki} = 0 \tag{3.2}$$

on each triple overlap $\mathcal{U}_i \cap \mathcal{U}_j \cap \mathcal{U}_k$. This system of fs is taken to be reduced *modulo* those f_{ij}s of the form of differences

$$h_i - h_j \tag{3.3}$$

where h_i is defined on $\mathcal{U}_i$ and h_j on $\mathcal{U}_j$. This means that the fs are really defining an element of the *first holomorphic sheaf cohomology group* of the space, with respect to the covering $\{\mathcal{U}_j\}$, (Morrow and Kodaira 1971; Wells 1979). To get rid of the dependence on the specific covering $\{\mathcal{U}_j\}$ we should take the direct limit with respect to finer and finer coverings. However, in practice, it is usually not necessary to take such a limit, where the sets $\mathcal{U}_i$ are taken to be appropriately 'convex' (Stein manifolds).

The passage from the twistor f_{ij} description to a space-time description can now be achieved by means of a *contour integral*. A point in $\mathbb{M}$, or else in its complexification $\mathbb{CM}$, is represented by a complex projective line in $\mathbb{PT}$, whose topology is S^2. This S^2 intersects the cover $\{\mathcal{U}_j\}$ in some system of open sets which cover S^2. In the simplest case there are just two (connected) sets in the cover $\{S^2 \cap \mathcal{U}_i\}$, which intersect in an annular region, and the contour integration is achieved using a contour running round this annulus (Fig. 3.1). In more complicated cases, this simple closed contour must be replaced by a suitable branched contour.

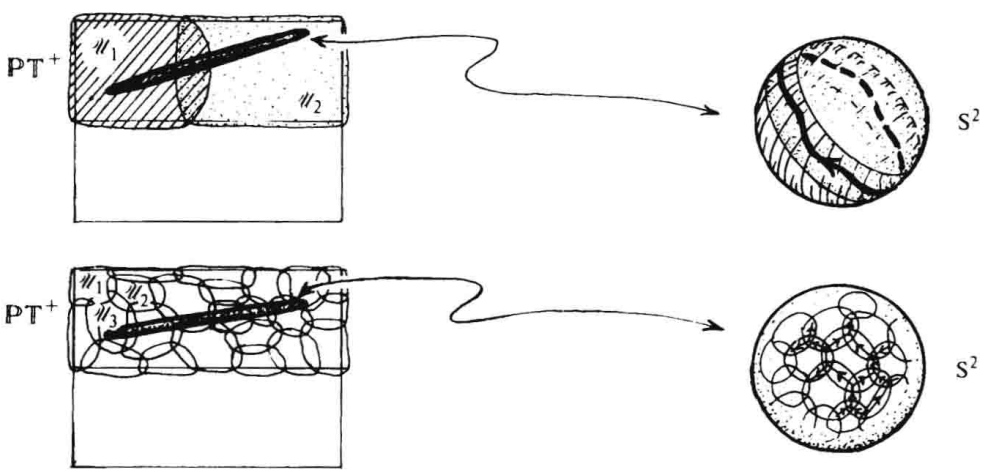


FIG. 3.1. The field at a space-time point is evaluated by contour integration.

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

The field defined in this way (Penrose 1969; 1979*a*,*b*; Penrose and Ward 1980; Ward this volume; Eastwood, Penrose, and Wells 1981) automatically satisfies the massless free field equations for helicity $\frac{1}{2}n\hbar$, where the twistor function f_{ij} is taken to be homogeneous of degree $-n-2$ in the twistor Z^{α}.

4. DEFORMED TWISTOR SPACES

A description in terms of first sheaf cohomology groups is very suggestive. For such cohomology groups may also play a role as *linearized deformations*. There is, indeed, a close similarity between functions defined on overlaps of pairs of open sets, subject to the relations (3.1) and (3.2), and *transition functions* used to build up a manifold out of coordinate patches. The key difference is that the relations (3.1) and (3.2) are linear, whereas the corresponding functional relations for transition functions are almost always non-linear. We may, indeed, obtain certain sheaf cohomology group elements as descriptions of *linearized* first-order deformations of manifolds. Then the 'gauge freedom' (3.3) arises from those linearized deformations which are trivial because they correspond simply to reparameterizations of the patches, defined by the h_is. In fact, finite (as opposed to infinitesimal) deformations can be thought of as being described by 'non-linear sheaf cohomology'. Then, in a sense, it is (3.3) which corresponds to the coordinate freedom in building a manifold, and (3.1) and (3.2) which correspond to the various compatibility relations for the transition functions.

This suggests that the sheaf cohomology group elements on twistor space that describe linear or non-interacting massless fields may perhaps be thought of as the linearized versions of some more general (non-linear) manifold construction. Indeed, this is precisely the case for twistor functions of homogeneity +2, describing linear massless particles of helicity $-2\hbar$ ($n = -4$): a non-linear version of such a twistor function gives rise to the 'non-linear graviton' construction whereby the general (if necessary, positive frequency) *anti-self-dual* solution of Einstein's vacuum equations is obtained from a deformed twistor space construction (Penrose 1976; Ward in this volume). Thus,

in a certain sense, we have 'one half' of (vacuum) general relativity precisely coded into twistor terms. That is, we can produce the general Ricci-flat 'space-time' with anti-self-dual Weyl curvature. To complete the picture we would have also to produce the *self-dual* part of the curvature by a twistor space construction. (Of course, the above term 'one-half' must be interpreted suitably liberally. One may note that the only space-times which are both anti-self-dual and Lorentzian are flat!)

The nature of the non-linear-graviton construction is such as to suggest, initially, that this programme is totally unpromising. At first sight it would seem that one ought to pass to the *dual* twistor space in order to have a non-trivial self-dual curvature, and to some kind of deformed *product* of the twistor space with its dual in order that *both* self-dual *and* anti-self-dual parts to the (Weyl) curvature may co-exist. However, this could not be viewed as a 'non-linear' version of the twistor function cohomology construction. For in the linearized theory, both the self-dual ($n = 4$) and anti-self-dual ($n = -4$) fields are described on the same *twistor* space, the dual space being not required. Of course the entire construction could be replaced by one in which the dual space $\mathbb{T}^*$ is used *throughout*, instead of $\mathbb{T}$. This makes no difference except to interchange the concepts of right- and left-handedness. But for a consistent quantum theory, one should not have to *change* the quantum description that is being used, merely because the helicity changes sign (which would be analogous to having to pass from a position to a momentum description whenever certain quantum numbers lie outside some range). Thus, whether we choose $\mathbb{T}$ or $\mathbb{T}^*$ we ought to be able to choose just one of them *consistently*, and to describe the solutions of the *full* Einstein vacuum equations, without having to resort to a hybrid description in which $\mathbb{T}$ and $\mathbb{T}^*$ are each used for part of the construction.

We recall now, that for the anti-self-dual case, a suitable non-linear version of the +2 homogeneity twistor function exists ($-n-2 = +2$, when $n = -4$). What we require next is a corresponding non-linear version of the -6 homogeneity twistor function ($-n-2 = -6$, when $n = 4$). To understand what is

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

presumably entailed, we think in terms of the (Poincaré invariant) exact sequence

$$\begin{array}{ccccccccc} 0 & \to & \mathbb{S} & \to & \mathbb{T} & \to & \tilde{\mathbb{S}}^* & \to & 0 \\ \omega^A & \mapsto & (\omega^A, 0) & & & & & & \\ & & (\omega^A, \pi_{A'}) & \mapsto & \pi_{A'} & & & & \end{array}$$

in the first instance, where $\mathbb{S}$ stands for the 'unprimed' spin space, and $\tilde{\mathbb{S}}^*$ for the dual 'primed' spin space. The standard non-linear graviton construction is a 'deformed' version of the interpretation of a space-time point as a cross-section of the fibration $\mathbb{T} \to \tilde{\mathbb{S}}^*$, i.e. a deformed version of the flat-space map

$$\begin{array}{ccc} & & \mathbb{T} \\ & & \uparrow \\ & & x \\ & & \downarrow \\ & & \tilde{\mathbb{S}}^* \end{array}$$

$$(ix^{AB'}\pi_{B'}, \pi_{A'}) \mapsto \pi_{A'} \qquad (4.1)$$

What we now require is a suitably 'deformed' version (referred to as the 'googly' graviton construction, Penrose 1979*a*,*b*) of the conjugate dual map

$$\begin{array}{ccc} & & \mathbb{S} \\ & & \uparrow \\ & & \tilde{x} \\ & & \downarrow \\ & & \mathbb{T} \end{array}$$

$$\omega^A - i\tilde{x}^{AB'}\pi_{B'} \mapsto (\omega^A, \pi_{A'}) \qquad (4.2)$$

The idea seems sound on very general grounds, but the detailed structure of the type of deformation that is required (which should depend on a non-linear version of a -6 homogeneity twistor function) seems remarkably elusive. At the present stage of understanding, there is no guarantee that such an idea will succeed. This is very much 'work in progress'. And it should be further remarked that even if it does succeed the problem of *combining* the deformed versions of (4.1) and (4.2) into a single scheme – wherein the most troublesome of the non-linearities of Einstein's equations would appear to reside (including those responsible for the violation of Huygens' principle, Ward in this volume) – would still remain to be solved.

In view of these most formidable difficulties, it may very reasonably be asked whether this programme has any chance at all of success. The non-linear graviton construction was found after all, in about 1975. If a solution to the remaining 'half' (or 'two-thirds') of the problem is to have any reasonable solution, then surely some clear indication of it should have become apparent in the five intervening years? To this I can only reply by pointing out some history. Twistor theory, as such, dates back essentially to 1964, while the 'twistor function' description of positive helicity (or self-dual) massless fields emerged in 1966. The description of negative helicity (or anti-self-dual) fields had to wait seven further years to appear! A simple suggestion made by Lane Hughston in about 1973 led to a new viewpoint in which the two helicities could be treated together within the same twistor scheme. Once having seen the suggestion, then, of course, all seemed obvious. But for seven years a somewhat misleading viewpoint had prevented us from noticing it! It should be pointed out also that it took us until 1976 to realize that we had been doing sheaf cohomology since 1966. (This was partly a matter of simple ignorance – though only partly. Michael Atiyah was very instrumental in educating us.) So it is, to me, not so surprising that we can as yet not see around whatever subtleties may be preventing us from understanding how general relativity is to be properly transcribed into twistor terms. In my own view, it is *necessary* for the theory that a solution can eventually be found. The theory ultimately stands or falls by it.

5. HYPERSURFACE TWISTORS

There are, however, various interesting things that twistor theory has to say about curved space-times, and about quantization in such backgrounds, even in the absence of a complete solution. One line of development concerns the construction of *hypersurface twistor space* (Penrose 1975). If we consider a normal Lorentzian space-time $\mathscr{M}$ (preferably globally hyperbolic) with a hypersurface $\mathscr{S}$ lying in it (preferably a spacelike Cauchy hypersurface – or else $\mathscr{I}^+$ or $\mathscr{I}^-$) then we can perform an analogous construction to that which produces $\mathbb{PN}$,

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

from $\mathbb{M}$, to obtain a manifold $\mathbb{P}\mathcal{N}$ as the space whose points represent null geodesics in $\mathcal{M}$. The local structure that $\mathbb{P}\mathcal{N}$ acquires – essentially the structure that arises when a 5-real-dimensional hypersurface resides in a 3-complex-dimensional ambient space (although here the structure can be defined intrinsically) – is called a CR-structure (Folland and Kohn 1972; Nirenberg 1973). In space-time geometrical terms, it describes the *shear* that systems of null geodesics possess, just as they encounter the hypersurface $\mathcal{S}$ (see Fig. 5.1).

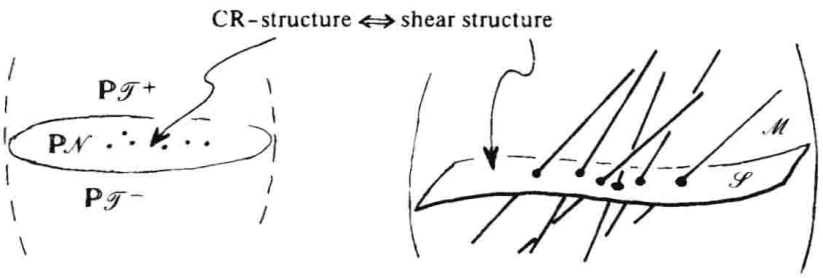


FIG. 5.1. The hypersurface twistor space for a hypersurface $\mathcal{S}$.

Assuming that $\mathcal{M}$ and $\mathcal{S}$ are analytic, it turns out that $\mathbb{P}\mathcal{N}$ is, indeed, a hypersurface in some ambient (projective) curved twistor space $\mathbb{P}\mathcal{T}$, which $\mathbb{P}\mathcal{N}$ divides into two portions $\mathbb{P}\mathcal{T}^+$ and $\mathbb{P}\mathcal{T}^-$. (In fact the space $\mathcal{T}$ can be interpreted as the space of solutions of the spinor ordinary differential equation $\pi^{C'}\pi_D, n^{BD'}\nabla_{BC'}\pi_{A'} = 0$, in the complexified space $\mathbb{C}\mathcal{S}$, n^a being the normal to $\mathbb{C}\mathcal{S}$ in $\mathbb{C}\mathcal{M}$. Factoring out by the equivalence $\lambda\pi_{A'} \equiv \pi_{A'}$ $(\lambda\pi_{A'} \neq 0)$ we obtain $\mathbb{P}\mathcal{T}$ (Penrose 1975).)

An essential difference, however, between the general construction and that which arises when $\mathcal{M}$ is (conformally) flat is that generally the CR-structure that $\mathbb{P}\mathcal{N}$ possesses (and therefore the very identities of points of $\mathbb{P}\mathcal{T}^{\pm}$) depends essentially on the location of the hypersurface $\mathcal{S}$. This is because, with conformal curvature, the Sachs equations (Sachs 1961; Penrose 1968) will entail a change in the shear as we pass along the various null geodesics. Thus, the CR-structure of $\mathbb{P}\mathcal{N}$ 'evolves with time' whenever $\mathcal{M}$ is conformally curved. (This is, strictly speaking, true even when $\mathcal{M}$ is static. In such cases, the CR-structure evolves *pointwise* on $\mathbb{P}\mathcal{N}$, as $\mathcal{S}$ is translated in time, even though as an entire abstract CR-manifold,

the structure of $\mathbb{P}\mathscr{N}$ remains unaltered with time.)

6. AN APPROACH TO CURVED-SPACE QUANTIZATION

For the moment, let us set aside the fact that the CR-structure of $\mathbb{P}\mathscr{N}$ evolves and consider, instead, what the CR-structure of $\mathbb{P}\mathscr{N}$ can do for us. It is known (Ward 1977) that initial data for a (suitable) massless field at $\mathscr{S}$ can be given in terms of the first holomorphic sheaf cohomology of (portions of) $\mathbb{P}\mathscr{T}$. By analogy with the situation with $\mathbb{PT}$, we may expect to have some sort of positive-/negative-frequency splitting, depending upon whether the cohomology refers to $\mathbb{P}\mathscr{T}^+$ or $\mathbb{P}\mathscr{T}^-$. To be able to define such a splitting is a key requirement for building up a quantum field theory.

Now there is a deep theorem (Andreotti and Hill 1972; Hill and MacKichan 1977) which can be used to relate the cohomology of (any portion of) $\mathbb{P}\mathscr{T}$ to a kind of cohomology (referred to as hyperfunction-$\bar{\partial}_B$-cohomology; see Folland and Kohn 1972; Nirenberg 1973) which resides entirely 'within' the hypersurface $\mathbb{P}\mathscr{N}$. Referring to this cohomology as $\mathcal{H}^p(\mathscr{Q})$ ($p = 0,1,\ldots$) and the corresponding holomorphic sheaf cohomology in $\mathbb{P}\mathscr{T}$ by $H^p(\mathscr{R})$, where $\mathscr{Q}$ and $\mathscr{R}$ are open subsets of $\mathbb{P}\mathscr{N}$ and of $\mathbb{P}\mathscr{T}$, respectively, the theorem implies

$$\mathcal{H}^1(\mathscr{U}\cap\mathbb{P}\mathscr{N}) = \frac{H^1(\mathscr{U}\cap\mathbb{P}\mathscr{T}^+)\oplus H^1(\mathscr{U}\cap\mathbb{P}\mathscr{T}^-)}{H^1(\mathscr{U}\cap\mathbb{P}\mathscr{T})} \qquad (6.1)$$

where $\mathscr{U}$ is any open set in $\mathbb{P}\mathscr{T}$. Now consider first the special case $\mathbb{P}\mathscr{T} = \mathbb{PT}$. Provided that the various regions in $\mathbb{P}\mathscr{T}$ occurring on the right-hand side of this equation are of the appropriate kind, namely (in essence) that they can be described as the union of a set of projective lines, then the H^1 cohomology groups on the right all refer to massless fields in the corresponding regions of $\mathbb{CM}$. (I am supposing that the homogeneity degree is fixed throughout, so we have massless fields of a specific fixed helicity.) If, in fact, $\mathscr{U}$ is the whole of $\mathbb{PT}$, then we have $H^1(\mathscr{U}\cap\mathbb{PT}) = H^1(\mathbb{PT}) = 0$, so (9) becomes

$$\mathcal{H}^1(\mathbb{PN}) = H^1(\mathbb{PT}^+)\oplus H^1(\mathbb{PT}^-)$$

Some remarks on twistors and curved-space quantization. *Quantum gravity 2: A second Oxford Symposium* Eds C. J. Isham, R. Penrose, and D. W. Sciama, 578–92. 1981

the two terms on the right describing, respectively, the positive- and negative-frequency fields. A similar splitting will occur in (6.1) in any case for which $H^1(\mathcal{U} \cap \mathbb{PT}) = 0$. George Sparling has suggested, in fact, that (6.1) be used as a basis for an approach to quantum field theory in curved space-time. He shows first, that the left-hand-side of (6.1) can be interpreted (for suitable regions $\mathcal{U}$, and in particular for a region $\mathcal{U}$ which covers the whole of $\mathbb{PN}$) as the space of initial data on the corresponding region of $\mathcal{S}$ for the appropriate massless fields (say Maxwell fields, so as to be safe from consistency conditions: Buchdahl 1958; Plebanski 1965). Thus, apart from the awkwardness of the 'denominator' in (6.1), this expression provides us with a *definition* of the required positive-/negative-frequency splitting.

We have two problems. First, we may have $H^1(\mathcal{U} \cap \mathbb{PT}) \neq 0$, so that the positive-/negative-frequency splitting is not quite uniquely defined. Second, the CR-structure of $\mathbb{PN}$ 'evolves with time' in the sense of being dependent on the choice of $\mathcal{S}$, so the positive-/negative-frequency splitting may be expected also to evolve with time. This is the phenomenon which is normally taken to be responsible for particle creation effects in curved space-time.

But, the first problem of the 'denominator' term in (6.1) still remains. One way to ensure the condition $H^1(\mathcal{U} \cap \mathbb{PT}) = 0$ is to require that $\mathcal{U}$ is a Stein manifold. But, Stein manifolds contain no projective lines, so now the interpretation in terms of fields in space-time necessarily breaks down. But, it must still be the case that $\mathcal{H}^1(\mathcal{U} \cap \mathbb{PN})$ has some space-time interpretation. We can choose arbitrarily small neighbourhoods of any point $\mathbf{p} \in \mathbb{PN}$ which are Stein, and the points of such a neighbourhood represent null geodesics in $\mathcal{M}$ lying close to the particular null geodesic p in $\mathcal{M}$ which corresponds to $\mathbf{p}$. Thus, whatever interpretation $\mathcal{H}^1(\mathcal{U} \cap \mathbb{PT})$ has in space-time, it must have to do with the propagation of fields in the neighbourhood of p (Penrose 1979*b*). In fact, $\mathcal{H}^1(\mathcal{U} \cap \mathbb{PN})$ is infinite-dimensional, even in the limit when $\mathcal{U}$ shrinks down to the point $\mathbf{p} \in \mathbb{PN}$ itself. This limiting infinite-dimensional space seems to correspond to propagation non-analytic information along p. There is a unique positive-/negative-frequency

splitting in this limit – as, indeed, there is whenever $\mathscr{U}$ is Stein. But the significance of this for curved-space quantum field theory is not at all clear. It is not even understood, at present, how the evolution in CR-structure that occurs when $\mathscr{S}$ is moved up in time affects this splitting. However, it should be a much easier question to examine than the general global question of the evolution of positive-/negative-frequency splitting since here we do not need to examine the structure of the entire hypersurface $\mathscr{S}$ and its embedding in $\mathscr{M}$, but merely at its intersection with the null geodesic p.

It would be interesting to examine the relation between Sparling's type of approach and the more conventional attacks on the problem of quantization (of massless fields) in a curved-space background. Clearly much more work needs to be done here.

One aspect of all this which I find intriguing is its possible relation to something that has always struck me as an irony inherent in the subject of quantum field theory. Classically, the concept of 'particle' arose out of attempts to localize the structure of matter into its smallest indivisible units. 'Locality' is, after all, central to the notion of a classical particle. Even in quantum field theory, a 'particle' ought intuitively to represent the most highly localizable of individual physical objects. Yet as theory stands, the concept of particle requires, strictly speaking, a knowledge of the structure of space-time even in its most distant regions. The detailed effect of the shape of distant regions can presumably be generally ignored. But it is an oddity that the concept 'particle' cannot really be *defined* in quantum field theory, without reference to them. Perhaps there may be new insights to be gained by use of an approach in which the concepts 'local' and non-local' have shifted their meanings!

REFERENCES

Andreotti, A. and Hill, C.D. (1972). *Ann. Scu. norm. sup. Pisa* CI. *Sci.* **26**, 325, 747.
Buchdahl, H.A. (1958). *Nuovo Cim.* **10**, 96.
Eastwood, M.G., Penrose, R., and Wells, R.O., Jr. (1981). *Commun. Math. Phys.* **78**, 305.

Folland, G.B. and Kohn, J.J. (1972). *The Neumann Problem for the Cauchy-Riemann Complex*. Princeton University Press.
Hill, C.D. and MacKichan, B. (1977). *Ann. Scu. norm. sup. Pisa* CI. *Sci.* **4**, 577.
Hughston, L.P. and Ward, R.S. (Eds) (1979). *Advances in Twistor Theory*. Pitman, London.
Morrow, J. and Kodaira, K. (1971). *Complex Manifolds*. Holt, Rienhard, and Winston, New York.
Nirenberg, L. (1973). *Lectures on Partial Differential Equations*. Amer. Math. Soc., Providence, Rhode Island.
Penrose, R. (1967). *J. Math. Phys.* **8**, 345.
—— (1968). In *Battelle Rencontres*. (ed. C.M. DeWitt and J.A. Wheeler). Benjamin, New York.
—— (1969). *J. Math. Phys.* **10**, 38.
—— (1975). In *Quantum Gravity, an Oxford Synposium*. (ed. C.J. Isham, R. Penrose, and D.W. Sciama). Oxford University Press.
—— (1976). *Gen. rel. Grav.* **7**, 31.
—— (1979*a*). In *Complex Manifold Techniques in Theoretical Physics*. (ed. D.E. Lerner and P.D. Sommers). Pitman, London.
—— (1979*b*). In *Advances in Twistor Theory*. (ed. L.P. Hughston and R.S. Ward). Pitman, London.
—— and MacCallum, M.A.H. (1972). *Phys. Reports*. 6, 241.
—— and Ward, R.S. (1980). In *General Relativity and Gravitation, One Hundred Years after the Birth of Albert Einstein*.Vol. 2. (ed. A. Held). Plenum, New York.
Plebanski, J. (1965). *Acta Phys. Polon.* **27**, 361.
Sachs, R. (1961). *Proc. R. Soc. Lond.* **A264**, 309.
Wells, R.O., Jr. (1979). *Differential Analysis on Complex Manifolds*. Springer-Verlag, New York.
Ward, R.S. (1977). *Curved Twistor Spaces*. D. Phil. thesis, Oxford.

123
ON SCHWARZSCHILD CAUSALITY—A PROBLEM FOR "LORENTZ COVARIANT" GENERAL RELATIVITY

Essays in general relativity (A. Taub Festschrift). Ed. F. J. Tipler, 1–12. Academic Press, 1980

A colleague has told me that he believes that there is an error in one of the detailed calculations presented here. I have unfortunately not had the time to check these calculations through again since then. I have no particular reason to doubt his contention, but I cannot believe that such an error would substantively alter the basic conclusions of the article, which are, at root, based on other clear geometrical arguments. Nevertheless, I would welcome clarification on this issue.

On Schwarzschild causality—A problem for "Lorentz covariant" general relativity. *Essays in General Relativity* (A. Taub Festschrift). Ed. F. J. Tipler, 1–12. 1980

1

On Schwarzschild Causality—A Problem for "Lorentz Covariant" General Relativity

Roger Penrose

Mathematical Institute
University of Oxford
Oxford, England

and

Department of Mathematics
University of California
Berkeley, California

I. Introduction

There appears to be a viewpoint, prevalent among some physicists [1] (cf. [2]), that while a geometrical approach to general relativity may have merits on aesthetic grounds and may have appeal for those whose interests are, perhaps, essentially pure mathematical, a strong emphasis on curved-space geometry is nevertheless to be rejected if real physical understanding and important future progress are to be achieved in gravitation theory. Abe Taub, however, is clearly not of this way of thinking. His many important contributions to both the physical and geometrical aspects of general relativity bear strong witness to the fact that far from being an obstacle to progress, differential geometry is an efficient and essentially indispensable tool in this highly significant aspect of physical insight.

As a way of honoring Abe's retirement, I shall present here a result that lends strong additional support to this geometrical viewpoint. It is directed, particularly, against the idea that general relativity might be adequately described as though it were a Lorentz-covariant (or, more correctly, Poincaré-covariant) field theory according to which the physical metric tensor is to be

treated as though it were not significantly different from any other field tensor.

II. Lorentz Covariance and Causality

Now the fundamental and unique role played by the physical metric tensor g (or at least by its conformal part $\hat{g}$, which represents 9 out of its 10 algebraically independent components) is, indeed, that it defines the *physical causal relations between points*. These causal relations play a key role in any classical relativistic field theory since they determine the propagation directions for *all* relativistic fields. Furthermore, the significance of this causal structure is as great in quantum field theory as in classical field theory. Quantum causality has the implication that field operators at spacelike-separated points must necessarily commute. If we were to take the standard "Lorentz-covariant" view, then we would need to introduce a background Minkowski metric η (perhaps not canonically) with the property that any two field operators at points that are spacelike-separated with respect to the *flat* causal structure defined by η would necessarily commute. (To modify this rule would be to reject the standard Lorentz-covariant viewpoint, from which all the standard results, such as the PCT and spin-statistics theorems are derived.) This is not to say that the final causality that is physically observed need agree with that defined by η. The actual way that fields propagate in the resulting theory would have to be calculated in detail. A normal procedure for doing this would be to obtain the metric g from a power series expansion of Lorentz-covariant terms, this being an infinite summation of Feynman graphs. (Summing "tree diagrams" is to give the classical g-field.) If such a Lorentz-covariant theory is to agree with general relativity, then the finally derived field propagation has to follow the null cones of this resulting general-relativistic curved metric g instead of those of η.

For a satisfactory theory, however, one would anticipate an important consistency requirement relating η to g: *the causality defined by g should not violate the background η-causality*. To put this another way, the g-null-cones ought never to extend outside the η-null-cones (Fig. 1). Thus, timelike curves with respect to g should always remain "timelike" with respect to η, i.e.,

$$g(T, T) > 0 \Rightarrow \eta(T, T) > 0 \qquad \text{at every point,} \tag{II.1}$$

for every tangent vector T (using the Lorentzian signature $+ - - -$). I write this condition

$$g < \eta. \tag{II.2}$$

 On Schwarzschild causality—A problem for "Lorentz covariant" general relativity. *Essays in General Relativity* (A. Taub Festschrift). Ed. F. J. Tipler, 1–12. 1980

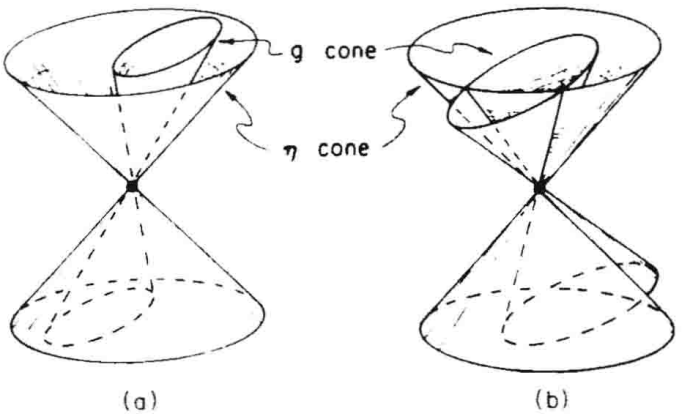


Fig. 1 Physical g-causality should not violate background η-causality. (a) Allowed: $g < \eta$ holds and (b) forbidden: $g < \eta$ fails.

If this condition were to *fail* to hold, then *physical* field propagation, which will follow the g-cones (or arbitrarily closely to them), would be superluminary with respect to η-causality. If we anticipate "physical" quantum field operators, describing the physically measurable fields, and constructible (say by infinite summation) from the original Lorentz-covariant operators, then these new field operators ought to be noncommuting for g-null- (or perhaps g-timelike-) separated points. But if $g < \eta$ fails, then these operators would have also to be noncommuting at certain pairs of η-spacelike-separated points (namely those along g-cones that extend outside η-cones). But this is not possible since noncommuting operators cannot be built out of the original commuting ones at spacelike η-separation.

Of course, in a proper quantized theory of general relativity, in which g also becomes a quantized field, the "g-cones" would never be perfectly well defined. However, if such a theory were to bear any resemblance to standard general relativity in the *classical limit*, then there should be a resulting "approximate" classical g-metric that would be anticipated to satisfy $g < \eta$, for the reasons outlined above. In the absence of a good theory of quantized gravity, there necessarily remains a certain inconclusiveness in this argument. Nevertheless, a question of some considerable interest for its *own* sake is whether or not, for a physically reasonable metric g on a space-time manifold M, a Minkowski metric η also exists on M with $g < \eta$. And if not, then this fact would also seem to provide some pertinent evidence against the fruitfulness of a Lorentz-covariant approach.

It is clear that for certain "strong gravitational fields" we should have difficulty in arranging $g < \eta$. For example, the *topology* of the space-time might itself differ from that of Minkowski space. Also there are space-times with causality violations (such as a maximally extended Kerr solution [3])

and certain examples (such as plane waves [4]), having no Cauchy hypersurfaces, and for which, for rather blatant reasons, $g < \eta$ cannot be arranged. But all these examples could be reasonably argued to be "nonphysical" and therefore not to constitute any significant case against the Lorentz-covariant viewpoint. However, I shall show here that the situation is much more serious than this. In even the simplest of physically reasonable nontrivial space–time geometries, namely, the Schwarzschild solution (with any suitable interior) $g < \eta$ indeed *cannot* be *satisfactorily* arranged.

III. Schwarzschild Geometry

At first, it would seem that the standard Schwarzschild coordinate system (t, r, θ, ϕ), for which the g-metric takes the form

$$ds^2 = \left(1 - \frac{2m}{r}\right)dt^2 - \left(1 - \frac{2m}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \quad \text{(III.1)}$$

provides a counter example to above contention, since all the null cones of this metric do in fact lie inside those of the corresponding Minkowski metric (in spherical polar coordinates):

$$d\sigma^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \quad \text{(III.2)}$$

provided that we restrict to the range

$$r \in [r_0, \infty) \quad \text{(III.3)}$$

for some $r_0 > 2m$. In order to cover the remaining range, we need an "interior" solution, such as the constant density perfect fluid solution originally proposed by Schwarzschild [1]

$$ds^2 = \left[\frac{3}{2}\left(1 - \frac{2m}{r_0}\right)^{1/2} - \frac{1}{2}\left(1 - \frac{2mr^2}{r_0^3}\right)^{1/2}\right]^2 dt^2 - \left[1 - \frac{2mr^2}{r_0^3}\right]^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \quad \text{(III.4)}$$

where

$$r \in [0, r_0], \qquad r_0 > 3m. \quad \text{(III.5)}$$

(I am not concerned with the question of black holes here, which would lead to some added complications.) These two solutions (III.1) and (III.4) match together to give a manifold with a C^0-metric g for which $g < \eta$, where the flat metric is given by (II.2). [It would not be hard, in fact, to modify the interior metric (III.4) to obtain a C^∞-metric g with this property, especially if we are not concerned with any particular equation of state in the interior

region.] We might think, therefore, that the effect of the Schwarzschild gravitational field is simply to "slow down" the velocity of light and so to provide $g < \eta$.

There is, however, an important sense in which this setup is quite unsatisfactory. This relates to behavior at infinity. One of the reasons for adopting a Lorentz-covariant viewpoint would, after all, be to discuss scattering theory. We might be concerned with incoming and outgoing asymptotically plane waves (e.g., gravitational perturbations or test Maxwell or massless scalar fields). Alternatively, taking the geometrical optics limit, we could be concerned with incoming and outgoing null geodesics. But, with regard to these null geodesics and null surfaces, the g-metric given by (III.1) actually differs very greatly from the η-metric given by (III.2) at large distances from the matter source (despite the fact that the metrics appear, naïvely, to go into one another in the limit $r \to \infty$). Consider, for example, the radial outgoing g-null geodesics given by

$$u = \text{const}, \qquad \theta = \text{const}, \qquad \phi = \text{const},$$

where u is g-retarded time:

$$u = t - r - 2m \log(r - 2m). \tag{III.6}$$

It is clear from the form of (III.6) that as $r \to \infty$, the value of $t - r$ is unbounded above along the geodesic. But $t - r$ is the standard η-retarded time, so we see that outgoing g-null geodesics reach indefinitely far into the η-retarded future, and do not correspond at all to outgoing η-null geodesics. Correspondingly, the incoming g-null geodesics are generators of the incoming g-null cones along which the g-advanced time

$$v = t + r + 2m \log(r - 2m) \tag{III.7}$$

is constant. Again there is no correspondence with the incoming η-null geodesics since the η-advanced time $t + r$ becomes unboundedly large and negative into the past along a g-null geodesic.

This is clearly unsatisfactory if this η-metric is to be used in a Lorentz-covariant approach for studying scattering theory in a Schwarzschild space time. On the other hand, we might envisage using an η-metric which is related to the Schwarzschild g-metric in a *different* way from the one just considered. To put things another way, we might choose a different coordinate system for the space-time, whose naturally associated flat metric was not the same as the η just considered. For example, we could use

$$r' = \tfrac{1}{2}(v - u) = r + 2m \log(r - 2m) \tag{III.8}$$

in place of r and now adopt the flat η'-metric given by

$$d\sigma'^2 = dt^2 - dr'^2 - r'^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{III.9}$$

In this particular case, since $u = t - r'$ and $v = t + r'$ are retarded and advanced times, respectively, for *both* metrics, the difficulty that was just mentioned concerning the difference in asymptotic structure between g and η does *not* now arise. However, it turns out that the condition $g < \eta$ is now what fails!

The purpose of this article is to show that there is, indeed, an essential incompatibility between the causal structure in Schwarzschild space-time (of positive mass m, and with *any* interior whatever) and that of Minkowski space. This shows up either asymptotically or in a violation of the local condition $g < \eta$. The most convenient way to handle the asymptotic conditions is by use of *conformal infinity* [3,5,6]. Both the Schwarzschild and Minkowski space-times have, in fact, well-behaved conformal infinities $\mathscr{I}^{\pm}$. Each point p of $\mathscr{I}^{+}$ can be thought of as describing an *outgoing (null) asymptotically plane wavefront*. (This is [3,7] the boundary $\partial I^{-}(p)$ of the TIP $I^{-}(p)$, cf. (IV. 2), representing p—an "event horizon" [6].) Likewise each point q of $\mathscr{I}^{-}$ describes an *incoming (null) asymptotically plane wavefront* (the boundary $\partial I^{+}(q)$ of the TIF $I^{+}(q)$ representing q—a "creation horizon" [3], sometimes referred to as a "particle horizon" [6]). Thus, $\mathscr{I}^{+}$ and $\mathscr{I}^{-}$ have very direct and natural interpretations in the context of scattering theory.

The space-time manifold with both conformal infinities $\mathscr{I}^{\pm}$ adjoined will be denoted $\overline{M}$. The conformal metric $\hat{g}$ is well-defined (C^0 will do) on the whole of $\overline{M}$, including its boundary $\mathscr{I}^{-} \cup \mathscr{I}^{+}$. The required condition that the asymptotic η-causal structure agree with the asymptotic g-causal structure (i.e., that the scattering theory be compatible for both metrics) is that both $\hat{g}$ and $\hat{\eta}$ be well defined on the *same* manifold-with-boundary $\overline{M}$. We have seen that this does *not* in fact hold for the $\hat{g}$ and $\hat{\eta}$ conformal metrics related to each other as is entailed by (III.1) and (III.2), but that it *does* hold for $\hat{g}$ and $\hat{\eta}$, related by (III.1) and (III.9).

IV. The Theorem

Let us now consider the more general question of whether or not a flat conformal $\hat{\eta}$-metric [not necessarily related to (III.1) by (III.2) or by (III.9)] exists at all on $\overline{M}$ and for which the required condition

$$\hat{g} < \hat{\eta} \tag{IV.1}$$

holds [this being the same as (II.2), but written now in terms of the conformal metrics, to emphasize that it applies to the *whole* of $\overline{M}$]. Here $\overline{M}$, with its given conformal metric $\hat{g}$, refers to the standard positive-mass exterior Schwarzschild solution [given by (III.1) in the range (III.3)] with any suitable

interior [such as (III.4) in the range (III.5)—but the choice of interior turns out to be irrelevant] and with standard conformal boundary $\mathscr{I}^+ \cup \mathscr{I}^-$ [with the points of $\mathscr{I}^+$ given by (u, θ, ϕ) at $r = \infty$ and those of $\mathscr{I}^-$ given by (v, θ, ϕ) at $r = \infty$]. The topology of $\overline{M}$ is built up from the (Schwarzschild) space-time:

$$M = \text{int } \overline{M} \cong \mathbb{R}^4$$

and its conformal boundary $\partial\overline{M} = \mathscr{I}^+ \cup \mathscr{I}^-$:

$$\mathscr{I}^+ \cong \mathscr{I}^- \cong S^2 \times \mathbb{R}.$$

With the interior (III.4) in the range (III.5), M is actually an asymptotically simple space-time [3,5,6]. The standard notation $a \ll b$, for points $a, b \in \overline{M}$ is adopted [6,8] for the assertion "there exists a future-directed timelike curve in $\overline{M}$ from a to b." Also

$$I^+(a) = \{x \in \overline{M} \,|\, a \ll x\}, \qquad I^-(a) = \{x \in \overline{M} \,|\, x \ll a\}. \tag{IV.2}$$

I shall prove:

(IV.3) *Theorem* $\mathscr{I}^+ \subset I^+(a)$ *for each* $a \in \mathscr{I}^-$; *equivalently,* $\mathscr{I}^- \subset I^-(a)$ *for each* $b \in \mathscr{I}^+$.

An equivalent statement but which does not refer to $\mathscr{I}^\pm$ is:

(IV.4) *Theorem* *If* λ *and* μ *are endless timelike curves in* M, *then there exist points* $p \in \lambda$, $q \in \mu$ *with* $p \ll q$.

From these results can be derived, as a simple corollary, the required property

(IV.5) *Theorem* *There is no Minkowskian conformal metric* $\hat{\eta}$ *on* $\overline{M}$ *with* $\hat{g} < \hat{\eta}$.

Proofs Let us first establish the equivalence of (IV.3) with (IV.4). Note that another way of stating (IV.3) is

$$a \ll b, \textit{ for every pair } a \in \mathscr{I}^-, b \in \mathscr{I}^+. \tag{IV.6}$$

Now, using the notion of causal boundary [3,7], we can interpret the point $a \in \mathscr{I}^-$ as an equivalence class of endless timelike curves having the same future in M (their points constituting the TIF in M representing a). Likewise, $b \in \mathscr{I}^+$ can be interpreted as an equivalence class of endless timelike curves with the same past in M (generating the TIP representing b). Let the endless timelike curves λ and μ represent a and b, respectively, so we can think of λ as acquiring a past endpoint a on $\mathscr{I}^-$ and μ as acquiring a future endpoint b on $\mathscr{I}^+$. The assertion $a \ll b$ amounts to saying that an endless timelike curve ν exists whose future agrees with that of λ and its past with that of μ. Now it is clear that *if* the statement of (IV.4) holds, then ν does indeed exist,

namely, consisting of the past-endless portion of λ up until the point p, the asserted timelike curve from p to q, and the future-endless portion of μ from q onward. (This jointed curve can be smoothed, if desired.) Conversely, suppose that λ and μ are as above and that ν has the same future as λ and the same past as μ. Choose a point w on ν. Now w must lie in the future of λ (since it lies in the future of ν), so a point p exists on λ with $p \ll w$. Similarly, q exists on μ with $w \ll q$. Hence $p \ll q$ holds, as required in (IV.4). There remains the possibility that one or both of the curves λ, μ in (IV.4) do not reach $\mathscr{I}^{\pm}$ at all, at their appropriate ideal endpoints but reach timelike infinity $i^{\pm}$ instead [3,5–7]. In these cases (IV.4) is again satisfied, but in a more trivial way.

Next, let us see why (IV.5) is a consequence of (IV.3) or (IV.4). Suppose a Minkowskian $\hat{\eta}$ exists on $\overline{M}$ with $\hat{g} < \hat{\eta}$. Now it is clear that (IV.4) would be false if "timelike" and "$\ll$" referred to the flat η metric. For example, in the usual T, X, Y, Z coordinates, with $d\sigma^2 = dT^2 - dX^2 - dY^2 - dZ^2$, the two branches of the η-timelike hyperbola $T^2 - X^2 + 1 = 0 = Y = Z$ are everywhere *η-spacelike* separated from one another. Suppose we let a be the past endpoint that one branch acquires on $\mathscr{I}^-$ and b the future endpoint that the other branch acquires on $\mathscr{I}^+$. By (IV.6), there is a g-timelike curve in $\overline{M}$ from a to b. But if $\hat{g} < \hat{\eta}$, this curve must also be η-timelike. But this is impossible as we have just seen. [The equivalence between (IV.4) and (IV.6) clearly holds equally for the $\hat{\eta}$-causality as for the $\hat{g}$-causality.] This contradiction establishes (IV.5) as a consequence of (IV.3) or (IV.4).

Finally, we must establish (IV.3). Let c be a point in M (with $r > r_0 > 3m$) and let γ be a null geodesic through c which is transverse to the source at c. Without loss of generality we can arrange c to have coordinates $(0, R, \pi/2, \pi/2)$ in the standard (t, r, θ, ϕ) coordinate system of (III.1) and, with the dot denoting $d/d\tau$, where τ is an affine parameter on γ,

$$\dot{r} = \dot{\theta} = 0 \qquad \text{at } t = 0. \tag{IV.7}$$

The standard geodesic equations for γ yield

$$\theta = \pi/2, \qquad r^2\dot{\phi} = A = \text{const.} \tag{IV.8}$$

and

$$[1 - (2m/r)]\dot{t} = B = \text{const.} \tag{IV.9}$$

We can normalize the affine parameter τ by choosing

$$B = 1 \qquad \text{and} \qquad \tau = 0 \qquad \text{at } t = 0. \tag{IV.10}$$

Then, by (IV.7)–(IV.10) (and taking the positive sign for $\dot{\phi}$),

$$A = R[1 - (2m/R)]^{-1/2}$$

and [using (III.6): $u = t - r - 2m \log(r - 2m)$], we derive

$$u = -R - 2m \log(R - 2m) + \int_R^r f(\rho)d\rho$$

where

$$f(\rho) = \left(1 - \frac{2m}{\rho}\right)^{-1}\left\{\left[1 - \frac{R^2}{\rho^2}\left(\frac{1 - 2m/\rho}{1 - 2m/R}\right)\right]^{-1/2} - 1\right\}.$$

By inspection of (IV.13) we see that

$$\int_R^\infty f(\rho)d\rho \tag{IV.11}$$

converges—as it must, since γ reaches a point of $\mathscr{I}^+$ with finite u-value. We wish to examine how this u-value behaves as $R \to \infty$ and to show, in fact, that it tends to $-\infty$:

$$\lim_{R\to\infty}\left[-R - 2m \log(R - 2m) + \int_R^\infty f(\rho)d\rho\right] = -\infty. \tag{IV.12}$$

Now, to see this, we note

$$f(\rho) < \left(1 - \frac{2m}{R}\right)^{-1}\left\{\left[1 - \frac{R^2}{\rho^2(1 - 2m/R)}\right]^{-1/2} - 1\right\} = g(\rho)$$

if

$$\rho \in (P, \infty) \qquad \text{where} \qquad P = R\left(1 - \frac{2m}{R}\right)^{-1/2} = R + O(1),$$

whence

$$\int_P^\infty f(\rho)d\rho < \int_P^\infty g(\rho)d\rho = R\left(1 - \frac{2m}{R}\right)^{-3/2} = R + O(1) \tag{IV.13}$$

by explicit integration. Furthermore we can estimate $\int_R^P f(\rho)d\rho$ using

$$\begin{aligned} f(\rho) &= \left(1 - \frac{2m}{\rho}\right)^{-1} \\ &\quad \times \left(\left\{\frac{(\rho - R)[\rho R(\rho + R) - 2m(\rho^2 + \rho R + R^2)]}{\rho^3(R - 2m)}\right\}^{-1/2} - 1\right) \\ &< \left(1 - \frac{2m}{R}\right)^{-1}(\rho - R)^{-1/2}P^{3/2}(R - 2m)^{1/2}(2R^3 - 6mP^2)^{-1/2}, \end{aligned}$$

provided that

$$\rho \in (R, P) \qquad \text{and} \qquad R > 5m$$

whence, using

$$\int_R^P (\rho - R)^{-1/2}\, d\rho = 2(P - R)^{1/2} = O(1),$$

we derive the fact that

$$\int_R^P f(\rho) d\rho = O(1) \tag{IV.14}$$

as $R \to \infty$. Combining (IV.14) with (IV.13), we obtain the result that (IV.11) is $R + O(1)$, so that substituting into the left-hand side of (IV.12), we derive the required result (IV.12), because of the presence of the logarithmic term.

This shows that whatever value of u is chosen, R can be made large enough that the null geodesic γ meets $\mathscr{I}^+$ at a u-value that is *less* than that chosen value. Moreover, because of the light-bending effect, γ will encounter an outgoing radial null geodesic β whose equation has the form

$$u = u_0 = \text{const}, \qquad \theta = \pi/2, \qquad \phi = \phi_0 = \text{const} \tag{IV.15}$$

for any value of ϕ_0 in the range

$$\phi_0 \in [\pi/2, \pi] \tag{IV.16}$$

(*including* the value $\phi_0 = \pi$), before reaching $\mathscr{I}^+$. Since u is an increasing function along γ, the value u_0 must be even less than the u-value attained at $\mathscr{I}^+$. We can likewise repeat the entire preceding argument in time-reversed form and attain the result that whatever value of v [given by (III.7)] is chosen, R can be made large enough that γ encounters, into the past, an incoming null geodesic α with equation

$$v = v_1 = \text{const}, \qquad \theta = \pi/2, \qquad \phi = \phi_1 = \text{const} \tag{IV.17}$$

for which v_1 is larger than the chosen v-value, and where ϕ_1 can take any value in the range

$$\phi_1 \in [0, \pi/2].$$

A jointed null geodesic, made up from pieces of α, γ, and β therefore connects the point $a \in \mathscr{I}^-$, with (v, θ, ϕ)-coordinates $(v_1, \pi/2, \phi_1)$, to the point $b \in \mathscr{I}^+$, with (u, θ, ϕ)-coordinates $(u_0, \pi/2, \phi_0)$. Thus [8]

$$a \ll b$$

as is required for (IV.6). (The jointed null geodesic can be smoothed, if desired, to yield to smooth timelike curve from a to b.) By a suitable rotation of the (θ, ϕ)-coordinate system, we can arrange for a to lie on any generator of $\mathscr{I}^-$ and b on any generator of $\mathscr{I}^+$. (The crucial case, in fact, is when the generators are opposite: $\phi_0 = \pi$, $\phi_1 = 0$.) And by allowing v_1 to be as

large and positive as desired, and u_0 to be as large and negative as desired, we can cover all possibilities, thus establishing (IV.3). ■

V. Concluding Remarks

It should be clear from the preceding construction that the results of this paper are in no way specific to the Schwarzschild solution. One is concerned only with the nature of the space-time at large distances from the positive-mass source—evidently essentially with causal properties in the neighborhood of spacelike infinity i^0 [3,5]. Corresponding results are to be anticipated for *any* appropriately asymptotically flat space-time with positive mass.

One is tempted to use the fact that, whenever the null convergence condition holds [3,9] (a consequence of the weak energy condition and Einstein's equations) together with the genericity condition [3,9], every complete null geodesic in the space-time contains pairs of conjugate points [3,9]. This has the implication that for any point $a \in \mathscr{I}^-$ (assuming asymptotic simplicity), no generator of $\partial I^+(a)$ in the space–time extends all the way to $\mathscr{I}^+$. This imposes severe difficulties for the geometry of $\partial I^+(a) \cap \mathscr{I}^+$, if this set is to be nonvacuous, and appears to lead to a more general argument from which (IV.3) can be derived under much wider circumstances. The question seems also to be related to the positive energy conjecture [10,11] and to the details of the structure [12] of i^0. These matters will not be discussed here, as the result obtained in Section IV is adequate for the present purposes.

Acknowledgment

The author is grateful to the Miller Foundation and the University of California at Berkeley for their support and hospitality.

References

[1] Weinberg, S., "Gravitation and Cosmology." Wiley, New York, 1972.
[2] Duff, M. J., *in* "Quantum Gravity." C. J. Isham, R. Penrose, and D. W. Sciama, eds., p. 78-135. Oxford Univ. Press, London and New York, 1975.
[3] Hawking, S. W., and Ellis, G. F. R., "The Large Scale Structure of Spacetime." Cambridge Univ. Press, London and New York, 1973.
[4] Penrose, R., *Rev. Mod. Phys.* **37**, 215 (1965).
[5] Penrose, R., *Proc. R. Soc. London, Ser. A* **284**, 159 (1965).
[6] Penrose, R., *in* "Battelle Rencontres" (C. M. DeWitt and J. A. Wheeler, eds.), p. 171–189. Benjamin, New York, 1968.

[7] Geroch, R. P., Kronheimer, E. H., and Penrose, R., *Proc. R. Soc. London, Ser. A* **327**, 545 (1972).
[8] Penrose, R., "Techniques of Differential Topology in Relatively." *SIAM*, Philadelphia, Pennsylvania, 1972.
[9] Hawking, S. W., and Penrose, R., *Proc. R. Soc. London, Ser. A* **314**, 529 (1970).
[10] Geroch, R. P., *Ann. N.Y. Acad. Sci.* **224**, 125 (1973).
[11] Schoen, R., and Yau, S.-T., *Commun. Math. Phys.* **65**, 45 (1979).
[12] Ashtekar, A., and Hansen, R. O., *J. Math. Phys.* **19**, 1542 (1978).

124
CONCERNING A FOURIER CONTOUR INTEGRAL

Twistor Newsletter **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

Concerning a Fourier Contour Integral. *Twistor Newsletter* **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.4 **Concerning a Fourier Contour Integral** *by R. Penrose* (TN 11, February 1981)

There are various places in twistor theory where, given a function $F(\dot{Z})$, holomorphic in some region of $\mathbb{T}$, one is required to consider a corresponding operator $F(\dot{\partial}_W)$ (where $\dot{\partial}_W = \frac{\partial}{\partial z}$). Moreover, it might be that we are more interested in some cohomology class that $F(\dot{Z})$ serves to define than in the function F itself. The definition of the corresponding "cohomology class" for $F(\dot{\partial}_W)$ would then be needed (cf., for example, Penrose §III.4.3). Now one approach to the study of expressions like $F(\dot{\partial}_W)$ might be to use Fourier transforms. But the normal Fourier transform of a function $\varphi(x)$ requires φ to be defined along the entire real line, and this is unsuitable for holomorphic sheaf cohomology. Suppose, instead, then we have $\varphi(\zeta)$ holomorphic in some annular region in $\mathbb{C}$ ($\ni \zeta$), and define

$$f(z) = \frac{1}{2\pi \mathrm{i}} \oint e^{\zeta z} \varphi(\zeta) d\zeta,$$

the contour being taken around the annulus. This provides an entire function f, while φ must have singularities within the contour in order that f be non-zero. Note that if

$$\varphi(\zeta) = \ldots + a_2\zeta^{-3} + a_1\zeta^{-2} + a_0\zeta^{-1} + a_{-1} + a_{-2}\zeta + \ldots,$$

then

$$f(z) = a_0 + a_1 z + a_2 \frac{z^2}{2!} + a_3 \frac{z^3}{3!} + \ldots$$

so that the information contained in a_{-1}, a_{-2}, a_{-3}, ... gets lost in the passage to f. Note also that if $\varphi \mapsto \zeta\varphi$ then $f \mapsto \frac{d}{dz} f$ and that if $\varphi \mapsto \frac{d}{d\zeta}\varphi$ then $f \mapsto -zf$. The operation $\varphi \mapsto \zeta^{-1}\varphi$ corresponds to $f \mapsto \int f dz$, the coefficient a_{-1} now emerging as the constant of integration. Likewise $\varphi \mapsto \zeta^{-2}\varphi$ corresponds to $\int(\int f dz)dz$ and a_{-1}, a_{-2} both play roles as constants of integration, etc. So we may think of the freedom in φ, in the passage to f, as referring to all the potential constants of integration that might arise. For an inverse transformation we can write

$$\varphi(\zeta) = \int_{P(\zeta)}^{\propto\infty} e^{-\zeta z} f(\zeta) d\zeta$$

where $P(\zeta)$ is holomorphic on and within the annulus, and where $\propto\infty$ refers to an approach to infinity along a direction chosen to make the integral converge (which

requires suitable growth restrictions on f). The freedom in $P(\zeta)$ relates to the freedom in $a_{-1}, a_{-2}, a_{-3}, \ldots$. Taking $P(\zeta) = 0$ yields $a_{-1} = a_{-2} = \ldots = 0$. It is possible to show directly by moving contours around (modulo holomorphic pieces) that the double transform $f \rightsquigarrow \varphi \rightsquigarrow f$ is the identity, a new feature (for twistor theorists) arising here being integration into an essential singularity (at ∞).

$f(z)$	$\varphi(\zeta)$
$z^n/n!$	ζ^{-n-1}
e^{az}	$(\zeta - a)^{-1}$
$J_0(\sqrt{a\zeta})$	$e^{a/\zeta}$
$\begin{pmatrix} \log z \\ e^{a/z} \\ H_0(\sqrt{az}) \end{pmatrix}$	$\begin{pmatrix} \zeta^{-1}(\gamma + \log \zeta) \\ H_0(\sqrt{az}) \\ E(a/2) \quad \text{(Euler)} \end{pmatrix}$

No doubt all this is very well-known to the right people — but I do not know who they are. And the relevance to cohomological $F(\partial_W)$ is yet obscure. Thanks to Andrew Hodges.

125
A NEW ANGLE ON THE GOOGLY GRAVITON

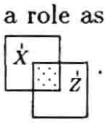
Twistor Newsletter **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

This is one of many approaches to the googly issue for gravitation. There does appear to be something 'correct' underlying this idea, but it is tantalizingly not *quite* right, even in the linear limit. These issues have never been properly resolved, as of the time of writing (May 2010).

A new angle on the googly graviton. *Twistor Newsletter* **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 264–9. 2001

§III.4.3 **A New Angle on the Googly Graviton** *by R. Penrose* (TN 11, February 1981)

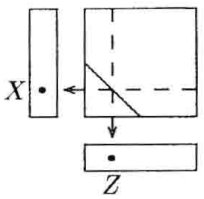
Recall, §O.3.9 and §III.4.2, the 'googly' idea for describing self-dual (as opposed to anti-self-dual) solutions of Einstein's vacuum equations in terms of some sort of deformed twistor space $\mathcal{T}$. Here $\mathcal{T}$ is 'flat' as a 'nonlinear graviton' (of the old-fashioned type) but now the information of the self-dual curvature is to be stored in the way that the line I sits in $\mathbb{P}\mathcal{T}$, or rather, in the way that $\mathbb{S}$ (the ω^A spin-space) sits in $\mathcal{T}$, in $0 \to \mathbb{S} \to \mathcal{T} \to \widetilde{\mathbb{S}}^* \to 0$ (in some sense). This is to be such as to determine a 4-parameter family of maps $x : \mathcal{T} \to \mathbb{S}$ (the 'googly maps') which generalize the flat space googly maps $x : (\omega^A, \pi_{A'}) \mapsto \omega^A - ix^{AA'}\pi_{A'}$, the regularity conditions on such maps being defined, at $\mathbb{S}$, by 'the way $\mathbb{S}$ sits in $\mathcal{T}$'.

Unfortunately, no good idea seemed forthcoming about how actually to fix these regularity conditions and, in particular, no relation had emerged for determining them, in the weak field limit, from a twistor function $f(Z^\alpha)$ homogeneous of degree -6. In the standard 'nonlinear graviton' a $+2$-homogeneous twistor function g plays a role as an infinitesimal transition function between two patches: $\dot{X} = \dot{Z} + \varepsilon \lfloor \partial_Z g(Z)$, 

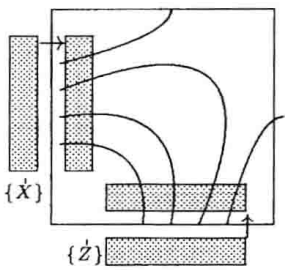
.

Our 'new angle' is to view f not as providing a transition function between patches, but somehow as providing a *co-transition function* between *co-patches!* The guiding idea is that we must reverse directions of maps (since the 'n.l.g.' setup $\mathcal{T} \underset{\tilde{x}}{\overrightarrow{\cdots}} \widetilde{\mathbb{S}}^*$ is to dualize

to the 'googly' setup $\mathbb{S} \overset{x}{\dashrightarrow} \mathcal{T}$). A 'patching' may be thought of in terms of a subspace of the product space of $\dot{X}$ with $\dot{Z}$:



and we need to use the two canonical projections $(\dot{X}, \dot{Z}) \to \dot{X}$ and $(\dot{X}, \dot{Z}) \to \dot{Z}$. So, dually, a 'co-patching' may be thought of in terms of a *factor space* of this product space:



where now we seem to be making use of the two canonical injections $\dot{X} \mapsto (\dot{X}, 0)$ and $\dot{Z} \mapsto (0, \dot{Z})$. To specify such a co-patching we require some sort of foliation of $\{\dot{X}\} \times \{\dot{Z}\}$, the 4-dimensional leaves of which integrate to the 4-surfaces that we factor out by to get $\mathcal{T}$. This might possibly relate to earlier googly ideas involving non-Hausdorff manifolds since such 4-surfaces might sometimes bifurcate.

How can we specify such a foliation, bearing in mind that in the weak field limit a -6-homogeneity function f must be involved? The answer seems to lie in the twistor quantization procedure $\bar{z}^{\dot{A}} \rightsquigarrow -\partial_Z$. Thus, instead of a dual-twistor transition relation $U = W + \varepsilon \eth_W \tilde{f}(W)$, we have $\partial_X = \partial_Z - \varepsilon \left[\eth Z, \tilde{f}(\partial_Z) \right]$. The reason for the Lie bracket arising on the right here is that, quantum mechanically, we would think of the $\eth_W$ of $\varepsilon \eth_W \tilde{f}(W)$ as acting on everything to the right of it, including some unmentioned ket $|\psi\rangle$ which is understood to be sitting over on the right, so that term should, instead, be written as $\varepsilon \left[\eth_W, \tilde{f}(W) \right]$, to make sure that only the $\tilde{f}$ is differentiated.

A function $\Phi(\dot{X}, \dot{Z})$ is deemed to be defined on the googly space $\mathcal{T}$ if it satisfies the equation $\partial_X \Phi = \partial_Z \Phi - \varepsilon \left[\eth Z, \tilde{f}(\partial_Z) \right] \Phi$, this being the statement that it is constant along the foliation. Now this is clearly a bit of a mess, because $\tilde{f}(W)$, which is the twistor transform of the -6-function $f(\dot{Z})$, is a $+2$-function which is nothing like a

A new angle on the googly graviton. *Twistor Newsletter* **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 264–9. 2001

polynomial, so $\tilde{f}(\partial_Z)$ is not simple to define. We can get some insight by considering the case of an elementary state. If we take

$$\tilde{f}(W_\alpha) = \frac{(W_0)^p (W_1)^q}{p!q!(W_2)^{r+1}(W_3)^{s+1}}, \qquad \text{then} \qquad f(Z^\alpha) = \frac{(Z^2)^r (Z^3)^s (-1)^{r+s}}{r!s!(Z^0)^{p+1}(Z^1)^{q+1}}$$

will do, where the coordinates Z^0,Z^1,Z^2,Z^3 and dual coordinates W_0,W_1,W_2,W_3 refer to some arbitrary basis (bearing no special relation to $I^{\alpha\beta}$). We need to interpret

$$\tilde{f}\left(\frac{\partial}{\partial Z^\alpha}\right) = \frac{\left(\frac{\partial}{\partial Z^0}\right)^p \left(\frac{\partial}{\partial Z^1}\right)^q}{p!q!\left(\frac{\partial}{\partial Z^2}\right)^{r+1}\left(\frac{\partial}{\partial Z^3}\right)^{s+1}}$$

and for this we can use

$$\int_0^u \left(\int_0^{u_1} \left(\cdots\left(\int_0^{u_r} \phi(u_{r+1})du_{r+1}\right)\cdots\right)du_2\right)du_1 = \int_0^u (-1)^r \frac{(x-u)^r}{r!}\phi(x)dx$$

(integrate repeatedly by parts) for the denominator terms and

$$\frac{1}{p!}\left(\frac{d}{du}\right)^p \phi(u) = \frac{1}{2\pi i}\oint \frac{\phi(x)}{(x-u)^{p+1}}dx$$

for the numerator terms. This gives

$$\tilde{f}\left(\frac{\partial}{\partial Z^\alpha}\right)\Psi(Z^\alpha) = \frac{1}{(2\pi i)^2}\oint\!\!\oint f(Q^\alpha - Z^\alpha)\Psi(Q^\alpha)d^4Q,$$

using some suitable contour-with-boundary in Q-space. This formula now makes no reference to the fact that we are using an elementary state, so we can extend by linearity to an arbitrary pair $\tilde{f}$, f which are twistor transforms of one another. The condition for Φ to be *defined on* $\mathcal{T}$ now becomes

$$\begin{aligned}(\partial_X - \partial_Z)\Phi(\dot{X},\dot{Z}) &= \frac{-\varepsilon}{(2\pi\,\mathrm{i})^2}\oint\!\!\oint (\lceil\!\!\lceil Z f(\dot{Q}-\dot{Z}) - \lceil\!\!\lceil Q f(\dot{Q}-\dot{Z}))\Phi(\dot{X},\dot{Q})d^4Q \\ &= \frac{\varepsilon}{(2\pi\,\mathrm{i})^2}\oint\!\!\oint \lceil\!\!\lceil R f(\dot{R})\Phi(\dot{X},\dot{Z}+\dot{R})d^4R\end{aligned}$$

where the substitution $\dot{R} = \dot{Q} - \dot{Z}$ has been made, and the contour boundary correspondingly moved.

Note that $\lfloor\!\!\lfloor \partial_X \Phi = \lfloor\!\!\lfloor \partial_Z \Phi$, which is to say that any such Φ is defined on *two-tailed twistor* space, i.e., the 6-dimensional space of pairs $(\dot{X},\dot{Z})$ factored out by the equivalence relation:

$$(\dot{X},\dot{Z}) \equiv (\dot{X} + \rho\lfloor\!\!\lfloor, \dot{Z} - \rho\lfloor\!\!\lfloor), \quad \rho \text{ arbitrary}.$$

Thus, a two-tailed twistor has two 'π-parts' but only one 'ω-part,' namely, the sum of the ω-parts of $\dot{X}$ and of $\dot{Z}$. This is the dual concept to that of a *two-headed twistor*, which is a pair $(\dot{X}, \dot{Z})$ subject to the restriction $\dot{X}\rceil = \dot{Z}\rceil$, so that there are two ω-parts now, but only one π-part. The standard (left-handed) nonlinear graviton arises as a subspace of two-headed twistor space, while the googly (right-handed) one arises as a factor space of two-tailed twistor space (at least in the simplest case when only two patches or co-patches are needed).

How do we define the googly maps? The motivating idea (which seems not to be quite right — although this is not absolutely clear) is as follows. We seek, for each googly point x, a map $\tau^A : \mathcal{T} \to \mathbb{S}^A$ which is the sum of a function σ of $\dot{X}$ and a function κ of $\dot{Z}$, i.e.,

$$\tau^A(\dot{X}, \dot{Z}) = \sigma^A(\dot{X}) + \kappa^A(\dot{Z})$$

(σ and κ each homogeneous of degree one). We require τ^A to satisfy the equation that ensures that it be defined on $\mathcal{T}$; also we require ('identity on I') $\sigma^A(X^B, 0) = X^A$, $\kappa^A(Z^B, 0) = Z^A$. These conditions ensure that, in the case $f = 0$, we get the standard flat space googly maps. Interpreting the spinor index as a 'hooked lower twistor index' we would obtain

$$\sigma^A(\dot{X}) \text{ '=' } -\sqcup\!\!\sqcap_X \qquad \text{and} \qquad \kappa^A(\dot{Z}) \text{ '=' } -\sqcup\!\!\sqcap_Z$$

in the flat space case, where $\sqcap = P_{\alpha\beta}$ represents the point defining the googly map, described in the standard way ($P_{\alpha\beta} = -P_{\beta\alpha}$, $P_{[\alpha\beta}P_{\gamma\delta]} = 0$, $P_{\alpha\beta}I^{\alpha\beta} = 2$). In the non-flat case we would have an additional part to σ^A and κ^A of order ε, given by a contour-with-boundary integral expression. This appears to be not quite consistent. However, a slight modification of the integral expression works!

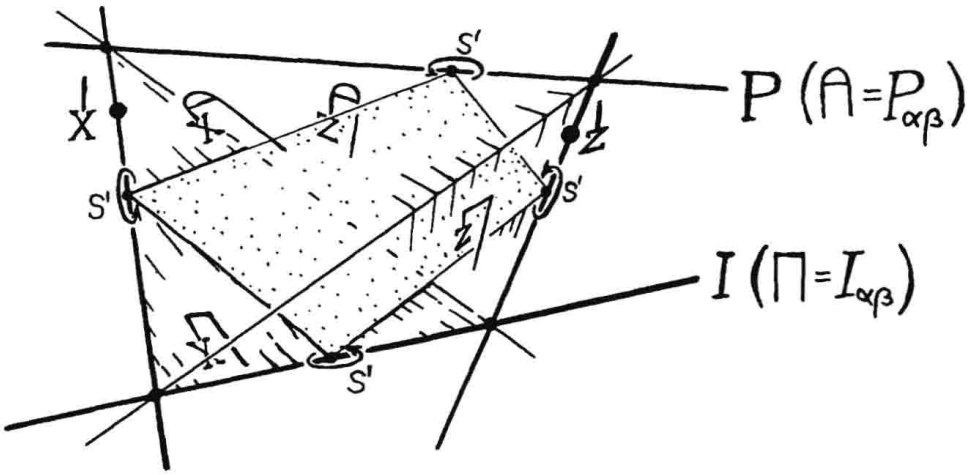
For this, we consider not quite σ^A, κ^A and τ^A, but slight generalizations $\varphi(\dot{X})$, $\chi(\dot{Z})$, $\gamma(\dot{X}, \dot{Z})$, and the googly maps are obtained by appropriately 'hooking' these expressions. We take $\varphi\{\dot{X}\} = \sqcap_X + O(\varepsilon)$, $\chi(\dot{Z}) = \sqcap_Z + O(\varepsilon)$ and $\gamma(\dot{X}, \dot{Z}) = \sqcap_X + \sqcap_Z + O(\varepsilon)$. Substituting τ for Φ in the above integral,

$$\partial_X\varphi(\dot{X}) - \partial_Z\chi(\dot{Z}) = \frac{\varepsilon}{(2\pi\,\mathrm{i})^2}\oint \lceil R \sqcap_R f(R) d^4R$$

is the expression that we come up with. We have to find an appropriate location for the boundary surfaces for the contour, anticipating (from the above) that the contour

A new angle on the googly graviton. *Twistor Newsletter* **11** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 264–9. 2001

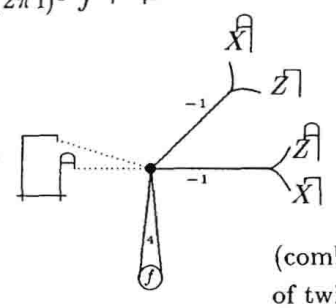
topology is, appropriately, $S^1 \times S^1 \times \mathbb{R} \times \mathbb{R}$ with a boundary probably looking like $\square \times S^1 \times S^1$ (i.e., $S^1 \times S^1 \times S^1$) (fill in the square for the contour itself).



What suggests itself is that the boundary lies successively in the planes $X\!\!\uparrow$, $Z\!\!\uparrow$, $Z\!\!\downarrow$, $X\!\!\downarrow$ and then back to $X\!\!\uparrow$ again, the contour being a quadrilateral area ($\times S^2 \times S^2$) spanning these planes. Then the integral yields a sum of a function of $\dot{X}$ and a function of $\dot{Z}$, as it must do, for consistency with the left-hand side, but unfortunately it fails the second test, which is that applying a further ∂_X on the left and skewing with the next 'leg' we should get zero. This problem is remedied by replacing our expression by its skew part:

$$\partial_X \sigma(\dot{X}) - \partial_Z \kappa(\dot{Z}) = \frac{\varepsilon}{(2\pi \mathrm{i})^2} \oint R\!\uparrow R\!\uparrow f(\dot{R}) d^4 R$$

$$= \varepsilon$$



(combining two types of twistor diagram!).

To obtain the googly maps we can evaluate $\partial_X \sigma \big|_{\substack{X \\ X}}$ or $\partial_Z \kappa \big|_{\substack{Z \\ Z}}$, (where the freedom in $\overset{\sigma}{X}$ may be supposed restricted by the condition $\overset{\sigma}{X} = 0$ (and $\overset{\kappa}{X} = 0$) *if desired*), σ (and κ)

being homogeneous of degree one (so $\partial_X \sigma = -\sigma$ '=' σ^A in *this* case).

Though the motivation for this expression has become somewhat muddy just at the end (!), one can check the formula in the case of the linearized Eguchi-Hanson elementary state, for which

$$f(\hat{Z}) = \frac{1}{\left(\begin{smallmatrix}A\\Z\end{smallmatrix}\right)^3 \left(\begin{smallmatrix}B\\Z\end{smallmatrix}\right)^3}$$

(cf. Burnett-Stuart 1979 and Penrose 1979b) and find (cf. also Hodges 1981) that

$$\partial_X \sigma = \frac{\left(B^{A^2\ B}_{X\ X} - A^{B^2\ A}_{X\ X}\right)^{A\ B}_{X\ X}}{\left(AB^{B\ A}_{X\ X}AB\right)^2} \qquad (\times \text{ numerical factor?})$$

which actually *is* the correct answer (as worked out by Law 1981)! Since *every* linear helicity +2 field is linearly dependent on such elementary states, this shows that the final formula works for a *general* f.

Note. In the above, I have used $\dot{\alpha} = -\sqcap = I^{\alpha\gamma}P_{\beta\gamma}$ ($= A^{\alpha}_{\beta}$ of Law's article) and $\psi = -\sqcup$, so $\dot{\alpha} = \sqcap$, $\dot{\psi} = \sqcap$, $\dot{\psi} = 0 = \dot{\sqcap}$, $\sqcap = 2$, $\dot{\psi} = \psi$, $\dot{\$} = 0 = \dot{\#}$, $\dot{\alpha} + \psi = 1$.)

Unfortunately it is not yet clear how to work all this when $\mathcal{T}$ is not merely infinitesimally co-deformed. The muddy logic at the end doesn't help matters. Further work is in progress. (Thanks to Paul Tod.)

References

Burnett-Stuart, G. (1979) Sparling-Tod Metric = Eguchi-Hanson, TN 9 (§III.1.3).

Hodges, A.P. (1981) Some More Elementary Twistor Integrals with Boundary, TN 11 (I.5.6).

Law, P. (1981) The Googly Maps for the Eguchi-Hanson/Sparling-Tod Graviton, TN 11 (§III.4.5).

Penrose, R. (1979a) A Googly Graviton?, TN 8, §O.3.9.

Penrose, R. (1979b) Remarks on the Sparling and Eguchi-Hanson (Googly?) Gravitons, TN 9, §III.4.2.

126
A QUADRATIC MAPPING WITH INVARIANT CUBIC CURVE

Math. Proc. Camb. Phil. Soc. **89** (1981), no. 1, 89–105. With C. A. B. Smith

Cedric Smith was a distinguished mathematical geneticist, who worked in my father's department (The Galton Laboratory) in University College London for a great many years, and was a good friend of the family. He died in 2002.

 A quadratic mapping with invariant cubic curve. *Math. Proc. Camb. Phil. Soc.* **89** (1981), no. 1, 89–105. With C. A. B. Smith

A quadratic mapping with invariant cubic curve

BY ROGER PENROSE
University of Oxford

AND CEDRIC A. B. SMITH
University College London

(*Received* 4 *September* 1979, *revised* 30 *April* 1980)

Abstract

In the projective plane, if H is a harmonic homology (linear transformation with $H^2 = I$), and G a general inversion (quadratic transformation projectively equivalent to an inversion), then under a certain condition there is a pencil of cubics each of which is invariant under G, H separately. These are related to transformations discovered by Mandel, Todd and Lyness. As a near converse, we find that, given a Pascal configuration, there is a quadratic Cremona transformation under which each cubic passing through the vertices of the configuration is invariant. As a by-product, parametric expressions are found for elliptic functions of a fifth of a period.

1. *Introduction*

The following investigation arose out of a classical problem in mathematical genetics. Let $\mathbf{x} = (x_0 : x_1 : \ldots : x_p)$ denote a point in p-dimensional projective space. Let $\mathbf{w} = (w_{rs})$ be a $p \times p$ matrix. Consider the mapping $M : \mathbf{x} \mapsto \mathbf{y} = \mathbf{x}^M$, defined by

$$y_r = \lambda x_r \sum_s w_{rs} x_s \quad (r, s = 0, 1, \ldots, p), \tag{1·1}$$

where λ is any non-zero constant, which without loss of generality may be taken to be 1. We are interested in the behaviour of the sequence of points obtained by iterating M, i.e.

$$\mathbf{x},\ \mathbf{x}^M,\ \mathbf{x}^{MM} (= \mathbf{x}^{M^2}),\ \mathbf{x}^{MMM},\ \ldots . \tag{1·2}$$

In the genetical application, $\mathbf{x}$ is interpreted as a set of gene frequencies in one generation, and w_{rs} the *fitness* or *viability* of the (rs) genotype. The vector $\mathbf{y} = \mathbf{x}^M$ then gives the gene frequencies in the next generation, $\mathbf{x}^{MM}$ the frequencies in the generation following that, and so on. In the genetical application we must have all $x_r > 0$ and all $w_{rs} = w_{sr} > 0$. This problem was investigated by Mulholland and Smith (10) and Scheuer and Mandel (13) independently, who discovered the main features of the behaviour of the sequence (1·2). See also (9). They showed that if the *mean fitness function* is defined as

$$\overline{w}(\mathbf{x}) = \Big(\sum_{r,s} w_{rs} x_r x_s\Big) \Big/ \Big(\sum_t x_t\Big)^2 \tag{1·3}$$

0305-0041/81/0000-7890 $03.50 © 1981 Cambridge Philosophical Society

then, apart from trivial exceptions, either

$$\mathbf{x} = \mathbf{x}^M = \mathbf{x}^{MM} = \ldots \quad \text{or else} \quad \overline{w}(\mathbf{x}) < \overline{w}(\mathbf{x}^M) < \overline{w}(\mathbf{x}^{MM}) \ldots \tag{1·4}$$

and the sequence (1·2) tends to a limit point $\mathbf{x}_\infty$ which is a local maximum of the function $\overline{w}(\ )$.

Much less is known about the behaviour of (1·2) when the restriction that $\mathbf{w}$ shall be symmetric and non-negative is removed. Mandel (personal communication) who investigated numerically for real values of the number a the particular case in which $p = 2$, and

$$\mathbf{w} = \begin{bmatrix} 1 & a & a^{-1} \\ a^{-1} & 1 & a \\ a & a^{-1} & 1 \end{bmatrix}, \tag{1·5}$$

found that the sequence of points (1·2) appeared to run cyclically round a closed curve in the plane. When the matrix has the particular form (1·5) we call the transformation (1·1) the Mandel mapping $\mathbf{y} = \mathbf{x}^{M(a)}$ (or, when a is taken as given, we write simply $\mathbf{y} = \mathbf{x}^M$).

Lyness (7, 8) remarked that any sequence obeying the recurrence relation

$$u_{n-1}u_{n+1} = 1 + u_n \tag{1·6}$$

is periodic, with period 5, as is readily verified algebraically. This can be written in a slightly different form due to H. Todd, as described by Lyness (8). Let $X^{(n)}$ be the point in projective 2-space with the coordinates

$$X^{(n)} = (1 : u_n : u_{n+1}).$$

Then the Todd–Lyness mapping $X \mapsto Y = X^L$ defined as

$$Y = (y_0 : y_1 : y_2) = (x_0x_1 : x_1x_2 : x_0[x_2 + x_0]) \tag{1·7}$$

transforms $X^{(n)}$ into $X^{(n+1)}$. It is a quadratic mapping with period 5. Various recurrences related to this have been known for a long time. Thus Burnside (2, p. 469) gives a formula for a birational transformation of period n. With n set equal to 5 this becomes

$$(\alpha_n, \alpha_{n+1}) \mapsto (\alpha_{n+1}, [-(1 + \alpha_n + \alpha_{n+1})/(1 + \alpha_n)]) = (\alpha_{n+1}, \alpha_{n+2}).$$

On replacing $-\alpha_n$, $-\alpha_{n+1}$ by x, y respectively, this becomes the transformation

$$(x, y) \mapsto (y, [1 - x - y]/[1 - x]), \tag{1·8}$$

mentioned by Baker (1, vol. 6, p. 121). If we set $u_n = -(1 + \alpha_n)$ it becomes Lyness's relation (1·6). Coxeter (4) has also related Lyness's relation to Napier's rules for spherical right-angled triangles, and to his own idea of frieze patterns.

This mapping L is a quadratic mapping, of a form rather similar to (1·1). We show in this paper that both Mandel and Todd–Lyness mappings are members of a family of quadratic mappings K in projective 2-space with the properties:

(1) given a mapping K, through any point X (with a finite number of exceptions) there passes exactly one cubic curve Γ such that K leaves Γ invariant;

(2) hence $X, X^K, X^{KK}, \ldots$, all lie on Γ:

(3) it is a well-known property (**1**, **5**, **15**) of plane cubics without a double point or cusp that they can be parametrized in such a way that the coordinates x_i are elliptic functions of the parameter t. The mapping K then simply adds a constant to the parameter:

$$t(x^K) = t(x) + \delta; \tag{1·9}$$

(4) hence the parameters of the points $X, X^K, X^{KK}, \ldots$ on the curve Γ are in arithmetic progression with common difference δ. In the particular case when δ is an exact submultiple of a period of the elliptic functions, the sequence $X, X^K, X^{KK}, \ldots$, is clearly periodic. For example, since the Todd–Lyness mapping has period 5, in it the corresponding δ must be $\frac{1}{5}$ of a period of the elliptic functions. In Section 7 of this paper this leads to explicit formulae for $\operatorname{cn}\delta$, $\operatorname{sn}\delta$, $\operatorname{dn}\delta$. More generally, when we consider real numbers and functions, it is not difficult to show that δ can be chosen to be real and the corresponding elliptic functions also have a real period. Hence, even if δ is not an exact submultiple of the period, the successive points $X, X^K, X^{KK}, \ldots$ will all fall on the curve Γ and run round it 'almost periodically'. This explains Mandel's findings.

2. *The Mandel mapping*

We first discuss in more detail the properties of the Mandel mapping, $M(a)$, or, written explicitly,

$$\left.\begin{aligned} y_0 &= x_0(x_0 + ax_1 + a^{-1}x_2),\\ y_1 &= x_1(x_1 + ax_2 + a^{-1}x_0),\\ y_2 &= x_2(x_2 + ax_0 + a^{-1}x_1). \end{aligned}\right\} \tag{2·1}$$

If $a = 1$ this is the identity mapping, while the particular values $0, \infty, -1, \exp(\pm 2\pi i/3)$ give mappings with respective periods 6, 6, 2, 3, the last being linear, simply multiplying the coordinates by cube roots of unity. It simplifies the language, without essential loss of generality, to think of the triangle of reference, $V_0V_1V_2$, as an equilateral triangle, as shown in Figure 1. (As a point of detail we note that, apart from a slight distortion for the sake of clarity, Fig. 1 shows the effect of successive transformations $M(3)$ on the point $(1:1:3)$. It will be seen that the twelfth transform is indistinguishable by eye from the starting value.) It is convenient to use from now on the convention that, if P_0, say, denotes a point $(p_0:p_1:p_2)$, then P_1, P_2 are respectively the points $(p_2:p_0:p_1)$ and $(p_1:p_2:p_0)$ got by cyclically permuting coordinates. Similar conventions apply to lines, etc.

The Mandel transformation $M(a)$ is $(1, 1)$ except for the fundamental points $C_0 = (0:a:-1)$, C_1 and C_2, which are respectively mapped onto the lines d_0, d_1, d_2, where d_0 is

$$d_0\colon x_0 + a^{-1}x_1 + ax_2 = 0,$$

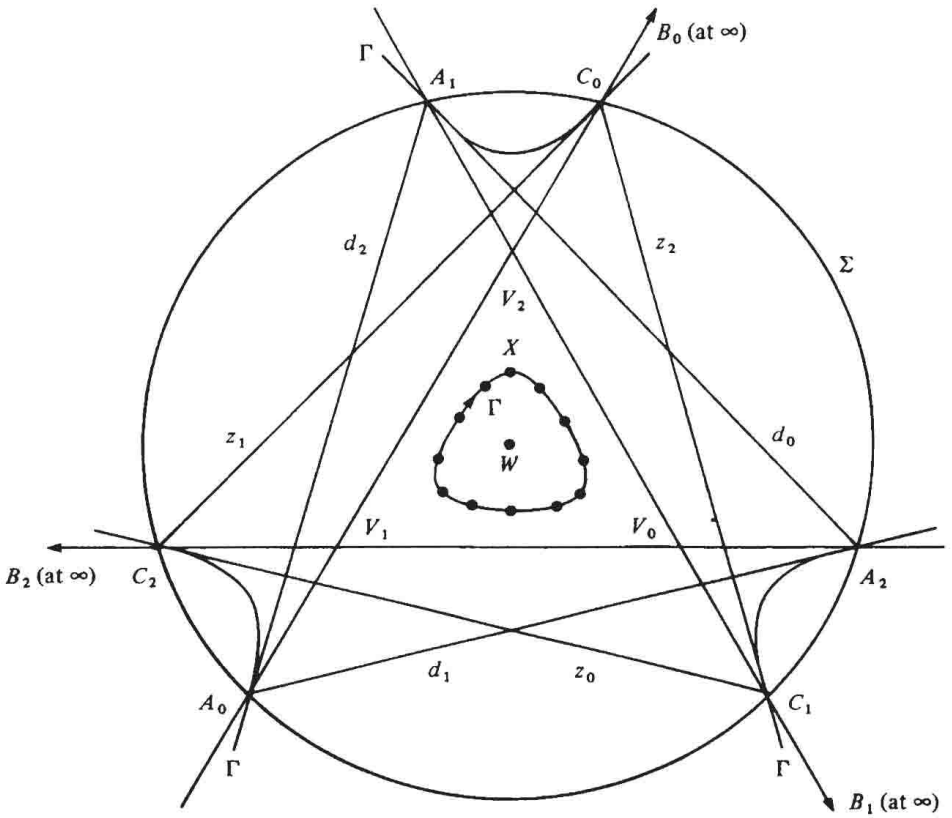


Fig. 1. Points, lines and curves related to the Mandel transformation $M = M(a)$. Γ is the invariant cubic curve. The points marked on Γ are successive transforms of an initial point X by the iterates of M. The diagram is slightly distorted for the sake of clarity.

and for points on the fundamental lines, z_0, z_1, z_2, where z_0 is

$$z_0\colon\ x_0 + ax_1 + a^{-1}x_2 = 0,$$

which are respectively mapped onto the points $A_0 = (0:1:-a)$, A_1 and A_2. In fact the Mandel transformation is a standard quadratic transformation (15) with respect to the net of conics through C_0, C_1, C_2; that is, it can be reduced to the form $(\xi_0:\xi_1:\xi_2) \mapsto (\xi_0^{-1}:\xi_1^{-1}:\xi_2^{-1})$ on taking $C_0C_1C_2$ as triangle of reference. Its fixed points are V_0, V_1, V_2 and $W = (1:1:1)$. If we also introduce the points $B_0 = (0:1:-1)$, B_1 and B_2 we see that $M(a)$ successively transforms z_i into A_i, A_i into B_i, B_i into C_i, and C_i into d_i (for $i = 0, 1, 2$). In the coordinate system used here, the line at $B_0B_1B_2$ has equation $x_0 + x_1 + x_2 = 0$, and transforms into the circle Σ with equation $\sigma(\mathbf{x}) = 0$, where

$$\begin{aligned}\sigma(\mathbf{x}) &= y_0 + y_1 + y_2\\ &= x_0^2 + x_1^2 + x_2^2 + (a + a^{-1})(x_0x_1 + x_1x_2 + x_2x_0).\end{aligned} \tag{2·2}$$

This circle passes through all A_i and all C_i, and is unaltered on replacing a by a^{-1}. Now we have identically from (2·1)

$$y_0 + a^{-1}y_1 + ay_2 = (a^{-1}x_0 + x_1 + ax_2)(ax_0 + a^{-1}x_1 + x_2),$$

so that if $M(a^{-1})$ transforms $Y = (y_0:y_1:y_2)$ into $X' = (x_0':x_1':x_2')$;

$$x_0' = y_0(y_0 + a^{-1}y_1 + ay_2) = (y_0y_1y_2/x_0x_1x_2)\,x_0 = \lambda x_0, \tag{2·3}$$

and x_1', x_2' follow cyclically, i.e. $X' = X$. Thus $M(a^{-1})$ transforms Y to X, and is the inverse of $M(a)$. Use of (2·2) with $M(a)$ replaced by $M(a^{-1})$ gives

$$\begin{aligned}\sigma(\mathbf{y}) &= x_0' + x_1' + x_2' \\ &= (y_0y_1y_2)\,(x_0 + x_1 + x_2)/(x_0x_1x_2),\end{aligned}$$

so that

$$\frac{(y_0 + y_1 + y_2)\,\sigma(\mathbf{y})}{y_0y_1y_2} = \frac{(x_0 + x_1 + x_2)\,\sigma(\mathbf{x})}{x_0x_1x_2} = c \quad \text{(say)}.$$

Hence if we define

$$\phi(\mathbf{x}) = (x_0 + x_1 + x_2)\,\sigma(\mathbf{x}), \tag{2·4}$$

we see that X and Y both lie on the cubic curve Γ_c with equation

$$\Gamma_c\colon \phi(x) = cx_0x_1x_2, \tag{2·5}$$

and hence, for any c, the cubic Γ_c is invariant under the transformation $M(a)$.

From now on, we confine our attention to one single value of a, and refer simply to the *Mandel transformation* M $(= M(a))$. For varying c, the Γ_c form a linear pencil of cubics. The 9 points A_r, B_r, C_r lie on all cubics of the pencil; through any other point there passes just one cubic. Degenerate members of the pencil include the following:

(i) when $c = \infty$, the 3 sides of the triangle of reference;
(ii) when $c = 0$, the line at infinity and the circle Σ;
(iii) when $c = (1-a)^3/a$, the triplet of lines $A_0B_1C_2$, $A_1B_2C_0$, $A_2B_0C_1$;
(iv) when $c = (a-1)^3/a^2$, the triplet of lines $A_0B_2C_1$, $A_1B_0C_2$, $A_2B_1C_0$.

There is a geometric construction for the mapping M (Fig. 2). Let R denote the reflection $X \mapsto X^R = (x_0:x_2:x_1) = Z$, say, so that $RR = I$, and let $Q = RM$ (operating from left to right, so that $Z^Q = (Z^R)^M = Z^{RM}$). Thus $M = RQ$, so that

$$Y = X^M = X^{RQ} = Z^Q.$$

By subtraction of coordinate vectors, $\mathbf{X} - \mathbf{Z} = (x_1 - x_2)\,\mathbf{B}_0$, so that X, Z and $F = B_0$ are collinear. Also we get at once from (2·1), which gives the coordinates of Y,

$$\mathbf{Y} - (x_0 + ax_1 + a^{-1}x_2)\,\mathbf{Z} = (x_1^2 + a^{-1}x_1x_0 - x_0x_2 - a^{-1}x_2^2)\,\mathbf{A}_0,$$

so that Z, Y and $E = A_0$ are collinear. Hence given X, we can construct Y by finding firstly the third intersection Z of the line XF with Γ_c, and then the third intersection Y of the line ZE with Γ_c.

It may be helpful to recall briefly and comment on some standard properties of cubic curves; see, for example (**1**, **5**, **15**) for details. Except where the contrary is explicitly stated, we allow a cubic to be degenerate (i.e. contain a straight line as a part), but we exclude from consideration degenerate cubics containing 2 or 3 coincident lines. Every non-degenerate cubic has at least one inflexion. By a suitable choice of coordinates, in which this inflexion is $(0:0:1)$, the equation of the cubic readily

reduces to the following form, most readily expressed in non-homogeneous coordinates,

$$y^2 = k(x-\alpha_1)(x-\alpha_2)(x-\alpha_3) = \psi(x) \quad \text{(say)}. \tag{2·6}$$

The Mandel cubic can be reduced to this form by taking

$$x_0:(x_1+x_2):(x_1-x_2) = (bx-1):(1+[c-b]x):y$$

where $b = (a-1)^2/a$. In the general case in which $\alpha_1, \alpha_2, \alpha_3$ are all distinct, we will say we have an elliptic (= non-singular plane) cubic curve; if two are equal, the cubic has a double point, and if all three are equal, a cusp.

Any non-degenerate cubic Γ can be parametrized by using the solution f of the differential equation

$$(f')^2 = \psi(f) \tag{2·7}$$

and setting

$$x = f(t), \quad y = f'(t). \tag{2·8}$$

In the 'elliptic' case, f is a doubly-periodic elliptic function: for a cubic with a double point, f is trigonometric (singly-periodic) and, for a cubic with a cusp, f is algebraic. In any case, by choosing a suitable f satisfying (2·7), we find that three points P_1, P_2, P_3 on the cubic Γ are collinear if and only if their parameters sum to zero (modulo the periods, if any, of f), i.e.

$$t(P_1)+t(P_2)+t(P_3) \equiv 0. \tag{2·9}$$

For the geometric construction given above (Fig. 2) for obtaining Y from X we have

$$t(Y)+t(Z)+t(E) \equiv 0,$$

$$t(X)+t(Z)+t(F) \equiv 0,$$

whence by subtraction

$$t(X^M)-t(X) \equiv t(Y)-t(X) \equiv t(F)-t(E) \equiv \delta \quad \text{(say)}, \tag{2·10}$$

which is a constant. Thus equation (2·10) holds for the Mandel mapping ($K = M$). In particular, (2·10) shows that

$$F = E^M \tag{2·11}$$

(as also readily follows from (2·1)).

3. *The Todd–Lyness mapping*

Similar arguments apply to the Todd–Lyness mapping, defined by (1·7), for which the inverse mapping is

$$Y \mapsto X = (x_0:x_1:x_2) = (y_0y_2:y_0[y_0:y_1]:y_1y_2), \tag{3·1}$$

which is also a quadratic Cremona transformation. The repeated transformation L^2 is also quadratic. If $Z = Y^L = X^{LL}$, then

$$(z_0:z_1:z_2) = (x_1x_2:x_2[x_0+x_2]:x_0[x_0+x_1+x_2]), \tag{3·2}$$

 A quadratic mapping with invariant cubic curve. *Math. Proc. Camb. Phil. Soc.* **89** (1981), no. 1, 89–105. With C. A. B. Smith

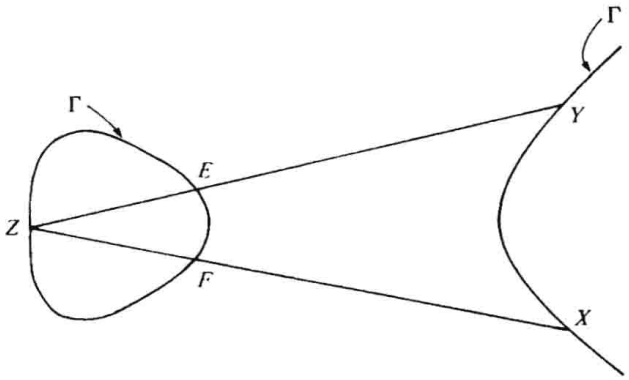


Fig. 2. Geometric construction for $Y = X^M$.

and similarly L^{-2} is quadratic. In fact for all integral n, L^n is either an identity mapping (when n is a multiple of 5) or else a quadratic Cremona transformation. If, following H. Todd [see (8)], we set

$$\phi_L(X) = (x_0+x_1+x_2)(x_0+x_1)(x_0+x_2), \tag{3·3}$$

and $Y = X^L$, then $\phi_L(X)/x_0x_1x_2 = \phi_L(Y)/y_0y_1y_2$, and hence by an argument like that in section 2 it again follows that the curves Γ_c^L,

$$\phi_L(X)/x_0x_1x_2 = \phi_L(Y)/y_0y_1y_2 = c \tag{3·4}$$

are invariant under L. The geometric construction (Fig. 2) for getting Y from X still holds, with the modification that, in defining Q, we take $E = (0:0:1)$ and $F = (0:1:-1)$. We can write L in the form $L = RQ$, where $Q = RL$, just as previously we had $M = RQ$. In either the Todd–Lyness or the Mandel case R is a reflection (and $R^2 = I$), Q is a quadratic mapping with $Q^2 = I$, and each of R and Q leave all the cubic curves Γ_c^L of the pencil (3·4) invariant.

4. *A family of quadratic mappings*

We now attempt to generalize these results to a whole family of mappings, which includes the Todd–Lyness and Mandel mappings as particular cases. Both L and M have the form RQ, where R is a (mirror) reflection and Q a quadratic transformation satisfying $Q^2 = I$. In either case, by a suitable change of coordinates from $(x_0:x_1:x_2)$ to, say, $(\xi_0:\xi_1:\xi_2)$, Q may be reduced to the mapping [as in (15), p. 47, ex. 3]:

$$(\xi_0:\xi_1:\xi_2) \mapsto (\xi_0^{-1}:\xi_2^{-1}:\xi_1^{-1}). \tag{4·1}$$

That is, Q is the product of a standard quadratic transformation

$$(\xi_0:\xi_1:\xi_2) \mapsto (\xi_0^{-1}:\xi_1^{-1}:\xi_2^{-1})$$

and an interchange of coordinates (ξ_1, ξ_2). Any (quadratic self-inverse) mapping reducible to the form (4·1) in a suitable coordinate system we call a *general inversion*, G. For the Mandel mapping it suffices to take $A_0A_1A_2$ to be the new triangle of reference (see Fig. 1). For the Todd–Lyness mapping set

$$\xi_0 = x_2, \quad \xi_1 = x_0, \quad \xi_2 = x_0+x_1. \tag{4·2}$$

A (linear self-inverse) mapping reducible to a reflection is a *harmonic homology* H. Thus, H has a *centre* point, C_H, an *axis* line λ_H not passing through C_H. The image P^H of a point H is the point on the line $C_H P$ such that P, P^H harmonically separate C_H and the intersection of $C_H P$ with λ_H. Thus $P^H = P$ if and only if either $P = C_H$ or P lies on λ_H. A mapping of the form

$$K = \Lambda G, \tag{4·3}$$

where Λ is linear and G a general inversion, is a *quadratic Cremona transformation*; the Mandel and Todd–Lyness mappings are particular cases.

It is thus natural to ask the question: given a harmonic homology H and a general inversion G, under what circumstances does there exist a cubic Γ invariant under H and G separately? To avoid detailed discussions of special cases, here and in what follows we will usually suppose that points, lines and other elements of our configurations are in 'general position'. What is meant by 'general position' will be defined more precisely when required.

If a cubic Γ invariant under a harmonic homology H does not contain the axis λ_H, then by taking C_H as the 0-vertex of the triangle of reference, and λ_H as the opposite side, the invariance under H (which simply changes the sign of the coordinate x_0) shows that the equation of Γ contains only even powers of x_0, from which it readily follows that C_H is an inflexion of Γ. Conversely, if C is any of the (in general 9) inflexions of Γ, there is a (unique) harmonic homology H with centre C leaving Γ invariant, and C lies on Γ. If Γ is invariant under the general inversion G it must pass through the fundamental points A_0, A_1, A_2 of G, for otherwise the transform of G would not be a cubic; see (15). Conversely, let Γ be a cubic (not 3 lines), let A_0 be a point on Γ not an inflexion, double point, or cusp, and not on a line forming part of Γ. Let the tangent to Γ at A_0 meet Γ again at N. Draw any line through N meeting Γ again at A_1, A_2, where A_0, A_1, A_2, N are all distinct. Then by taking suitable coordinates relative to the triangle $A_0A_1A_2$, we find that there is a general inversion of form (4·1) leaving Γ invariant.

Now suppose Γ invariant under H, G separately. Suppose further that H and G are in 'general position' in the sense that the C_H does not lie on any side of A_0, A_1, A_2; λ_H does not pass through any vertex A_0, A_1, A_2; $C_H^G \neq C_H$ and Γ does not contain the line A_0C_H. The points A_0^H and C_H lie on Γ and on the line A_0C_H, but are different from A_0 and C_H. Since this line can intersect Γ in only 3 points, the following condition must be satisfied,

$$A_0^H = C_H^G. \tag{4·4}$$

There is a stronger converse; if (4·4) holds then there exists a whole linear pencil of cubics invariant under H and G separately. To prove this we need two standard results (which we quote without proof).

Lemma 4·5. If P_1 to P_8 are 8 points such that no four lie on a line and no seven on a conic, then there is a (unique) linear pencil of cubics through P_r ($r = 1$ to 8) which also all pass through some ninth point P_9. (The points P_1 to P_9 are a set of 9 *associated points*.)

Lemma 4·6. If a conic passes through A_1, A_2 and 2 points of the form P, P^G, with all four points distinct, then it is invariant under G. (This is the projective form of the theorem that, if P, P^G are related by an inversion G in a circle, then any circle through P, P^G is invariant under G.)

As we have already seen, a cubic Γ (if there is one) invariant under H, G separately must pass through C_H, A_0, A_1, A_2, and therefore also through A_0^H, A_1^H, A_2^H, which together form a set Σ of 7 points. Conversely, let a cubic Γ pass through all 7 points of Σ; we proceed to show that Γ is invariant under H, i.e. if X is a general point of Γ, then X^H lies on Γ. By Lemma (4·5) there exists a ninth point Y such that all cubics through Σ and X (including in particular the cubic Γ) also pass through Y. But one such cubic Γ_1 consists of the line $A_1 A_1^H C_H$ together with the conic S_1 through $A_0 A_0^H A_2 A_2^H X$. Clearly C_H and λ_H are pole and polar with respect to S_1, so that S_1 is invariant under H. Hence S_1 passes through X^H, and so therefore does the cubic $\Gamma_1 = S_1 \cup$ line $A_1 A^H C_H$. By interchanging suffixes '1' and '2' we get a second cubic Γ_2 passing through X^H. But Y must be the intersection of Γ_1 and Γ_2 not contained in $\Sigma \cup \{X\}$; hence $Y = X^H$, i.e. every cubic through Σ and X goes through X^H. This shows that every cubic through Σ is invariant under H.

Suppose further that X is chosen so that X, X^G and X^H are all distinct. Then Σ, X, X^G form a set of 9 non-associated points, so there is a unique cubic Γ_X through them and by the preceding result it is invariant under H. Let Z be another point on Γ_X, such that $\Sigma \cup \{X, X^G, Z\}$ does not contain 4 points on a line or 7 on a conic. Consider the set Σ' of 8 points $\{A_0 A_1 A_2 A_0^H (= C_H^G) X X^G Z\}$: this has a ninth associated point Z'. But since Z' must lie on the 2 cubics

$$\text{line } A_0 A_0^H C_H \cup \text{conic } A_1 A_2 X X^G Z,$$

$$\text{line } A_0 X X^G \cup \text{conic } A_1 A_2 C_H C_H^G Z,$$

it must be their intersection not in Σ'. But Lemma 4·6 shows that both these conics are invariant under G; so also are the lines which pass through A_0, and hence both cubics are invariant under G and pass through Z^G, i.e. $Z' = Z^G$. Hence if Z lies on Γ_X so also does Z^G, i.e. Γ_X is invariant under G. Thus (provided that (4·4) is satisfied) through a general point X there passes just one cubic Γ_X invariant under H and G separately, i.e. such cubics form a one-dimensional family. They form a linear pencil because the invariances under H and G impose linear constraints on the coefficients: an alternative proof of this is given later.

It is natural to ask what will happen if instead of a harmonic homology H and a general inversion G we consider curves invariant under two harmonic homologies H, h, or under two general inversions G, g.

In the case of two harmonic homologies, H, h, it is not difficult to see that, if X, X^H, X^{Hh}, X^{HhH}, X^{HhHh} are all distinct, any conic through them (in general unique) is invariant under H and h separately. (Think of the images of a point placed before two mirrors.) This is equivalent to a theorem, discovered independently by Coxeter (3) and Komissaruk (6), that the orbit of any point for an equiaffinity lies on a conic. (N.B.

As regards cubics, we have already observed that an elliptic cubic is invariant in general under 9 harmonic homologies.)

In the case of 2 general inversions, G, g, let the corresponding triangles of reference (in general position) be $A_0A_1A_2$ and $a_0a_1a_2$. Any cubic Γ invariant under G and g will have to meet the line A_0a_0 in 4 points, A_0, a_0, A_0^g, a_0^G. If these are all distinct then the line A_0a_0 forms part of the cubic; the remaining part will be a conic invariant under both G and g. Such a conic will exist only in very special cases, and Γ will then be degenerate. Otherwise, at least 2 of A_0, a_0, A_0^g, a_0^G must coincide. If we consider the 'general' case, by which we mean here that no vertex of one of the triangles $A_0A_1A_2$, $a_0a_1a_2$ lies on a side of the other, then the only possible coincidence will be

$$A_0^g = a_0^G. \tag{4·7}$$

Any cubic invariant under G and g must then pass through the 9 points A_i, a_i, $A_0^g = a_0^G$, A_1^g, a_1^G. Apart from special cases in which 7 of them lie on a conic, or in which $A_1^g = a_1^G$, there will be just *one* cubic Γ through them. Furthermore, this cubic Γ is invariant under both G and g. For let Z be any point differing from 9 points above, and such that $Z^G \neq Z$. By Lemma (4·6) the conic through $A_1A_2a_0a_0^GZ$ is invariant under G, and hence the cubic consisting of the conic and the line $A_0a_1a_1^G$ is invariant under G, and must also pass through Z^G. Similarly the conic $A_1A_2a_1a_1^GZ$ and line $A_0a_0a_0^G$ also form a cubic passing through Z^G. The intersection of these 2 cubics is therefore a set of 9 associated points, i.e. any cubic through $A_0A_1A_2a_0a_0^Ga_1a_1^GZ$ passes through Z^G, i.e. any cubic such as Γ through $A_0A_1A_2a_0a_0^Ga_1a_1^G$ is invariant under G. By interchanging $A_0A_1A_2$ with $a_0a_1a_2$ it similarly follows that Γ is invariant under g.

5. *Transformations leaving a given cubic invariant*

We saw above that under certain conditions a cubic Γ is invariant under a harmonic homology H and a general inversion G separately and therefore also under their products GH, HG, GHG, etc. These transformations are all (1, 1) except at most on a finite number of lines. If Γ is in general position, in the sense that it does not contain any of these lines, then each of these transformations is analytic except possibly at a finite number of intersections of Γ with the exceptional lines.

We therefore consider the following question: if K is an analytic transformation of the plane inducing a transformation k on a nondegenerate cubic Γ, and if k is (1, 1) except possibly for a finite set of points (where k or its inverse may not be uniquely defined), what consequences follow? Let $t(X)$ be the parameter of the point X on Γ. Then a slight modification of a standard result (Segre, (14) pp. 735–6) shows that the transformation k is linear in the parameter:

$$t(X^k) = \alpha + \beta t(X). \tag{5·1}$$

Since our assumptions differ slightly from Segre's, we sketch a proof. Suppose first that Γ is elliptic, and $t(X)$ has periods $2\omega_1, 2\omega_2$. Convert the t-plane into a torus T^* by taking values modulo $2\omega_1, 2\omega_2$. The analytic transformation $X \mapsto X^k$ induces on T^* an analytic transformation κ which is (1, 1) except possibly on a finite set f^* of points, i.e., on a 'punctured torus'. We 'mend the punctures' by assigning a value of κ (or

κ^{-1}) at each 'puncture' p to make κ continuous, and hence necessarily analytic, at p, as follows. Choose one such (isolated) $p \in f^*$ with p^κ not yet uniquely defined. We want to show that, whatever sequence $\{t_n\}$ we choose tending to p as $n \to \infty$, we always get one and the same value L of $\lim t_n$, so that setting $p^\kappa = L$ ensures continuity. If that was not so, we could find some $\{t_n\} \to p$ such that $\{\tau_n\} = \{t_n^\kappa\}$ has more than one limit point, say $L_1, L_2, \ldots$. We may suppose $L_i \in f^*$, for otherwise κ^{-1} is analytic near L_i and must take the value p at L_i, mending the puncture. Separate L_1 from L_2 by a closed curve c not meeting f^*. Draw a continuous curve c' through $t_1, t_2, t_3, \ldots$ tending to p. Its transform c'^κ must meet c in an infinity of points, having some limit point L_0. Hence there is a sequence $\{t_n'\}$ on c' such that $\{t_n'\} \to p, \{t_n'^\kappa\} \to L_0$, a contradiction since $L_0 \notin f^*$.

After all punctures are mended, κ becomes a $(1, 1)$ analytic transformation of the torus T^* onto itself, readily converting to a $(1, 1)$ analytic transformation of the complex plane onto itself, which must be linear, justifying (5·1). Similar proofs apply when Γ is not elliptic.

From now on we restrict attention to elliptic Γ. Then the linear transformation (5·1) must map the lattice of periods onto itself. It is not difficult to see that in general this implies that $\beta = \pm 1$. Exceptions can occur only with the lattice of squares ($\beta^4 = 1$, $\beta = \pm i$) or of equilateral triangles ($\beta^6 = 1$). Such 'singular' transformations are dealt with in detail by Segre (**14**) and Zeuthen (**16**, pp. 223–237). We specifically suppose them excluded from our consideration.

Consider first $\beta = -1$, when (5·1) can be written

$$t(X) + t(X^k) + t(A) = 0, \tag{5·2}$$

where A is chosen so that $t(A) = -\alpha$. (Note that this excludes the possibility that A could be a double point or cusp, where the parameter is infinite.) Hence X^k is the third intersection of XA with Γ.

If A is an inflexion, we have seen that there is a harmonic homology H such that $X^k = X^H$ for all X on the curve. Hence H is the unique linear mapping of the plane such that $X^k = X$ for all X on Γ. Also there is no quadratic mapping Q such that $X^Q = X^k = X^H$ for all X on Γ. For H transforms any 3 collinear points on Γ into 3 collinear points, whereas Q in general would not.

If A is not an inflexion of Γ then we sketched in Section 4 a proof that there are ∞^1 general inversions G with centre A leaving Γ invariant. For X on Γ, X^G is the third intersection of XA with Γ, i.e. X^k. So any such G induces k on Γ. (Note, by the way, that these are not the only quadratic transformations with this property, or even the only self-inverse such quadratic transformations. Let Π be the polar conic of A with respect to Γ. Let X^F be the point on XA which together with X harmonically separates A and the other intersection of XA with Π. Then the mapping $F: X \mapsto X^F$ is self-inverse, quadratic, and induces k on Γ.)

Now consider the second case, $\beta = 1$, in (5·1), i.e.

$$t(X^k) = t(X) + \alpha. \tag{5·3}$$

Any inflexion A of Γ determines a harmonic homology H with centre A such that

$$t(X^H)+t(X)+t(A)\equiv 0,$$

so that

$$t(X^k)+t(X^H)+[t(A)-\alpha]\equiv 0.$$

Hence, by the previous argument, there exists either a second harmonic homology H' or a general inversion G mapping X^H onto X^k. That is, either there is a linear transformation HH' or a quadratic Cremona transformation HG inducing the mapping k on Γ. Also Γ is invariant under H alone, and also under H' or G alone as appropriate.

6. *The Pascal configuration*

We now consider the last case of a mapping of the form HG in more detail. In slightly greater generality, we suppose that a cubic Γ is invariant under a quadratic Cremona transformation, i.e. one of the form $K=\Lambda G$, where Λ is linear (but not necessarily a harmonic homology) and G is a generalized inversion, and that ΛG induces a transformation k of the form (5·3) on Γ. What follows? Note that we do not suppose here that Γ is necessarily invariant under Λ or G on its own. We write $\Lambda^{-1}=V$, so that $K^{-1}=GV$.

Let $A_0A_1A_2$ be the fundamental triangle of the general inversion G. We take suffixes modulo 3, so that $A_{-1}=A_2$, etc. Note that, if we take coordinates with respect to this triangle $A_0A_1A_2$, equation (4·1) defining G can be written $(\xi_i)\mapsto(\xi_{-i}^{-1})$ and this in turn implies that the line $A_{i-1}A_{i+1}$ is transformed by G into the vertex A_{-i}. Note also that, since $K^{-1}=GV$ transforms Γ into itself, Γ^G must be a cubic, and for this to be possible Γ must pass through all fundamental points A_i. Let A_i' be the third point of intersection of $A_{i-1}A_{i+1}$ with Γ. Provided that A_i' is distinct from A_{i-1} and A_{i+1} we have that

$$(A_i')^{K^{-1}}=(A_i')^{GV}=A_{-i}^V=C_i \quad \text{(say)} \tag{6·1}$$

is on Γ. If $A_i'=A_{i-1}$ or A_{i+1}, the line $A_{i-1}A_{i+1}$ touches Γ at A_i', and (for example) a consideration of the limit of X^{GV} as X on Γ tends to A_i' shows that $C_i=A_{-1}^V$ still lies on Γ. The line $C_{i-1}C_{i+1}$ meets Γ in a third point, C_i', say. Hence $C_i'^{\Lambda}$ lies on the line $C_{i-1}^{\Lambda}C_{i+1}^{\Lambda}$, and must be A'_{-i}, so that $C_i'^K=C_i'^{\Lambda G}=A_i$. (If C_i' coincides with C_{i-1} or C_{i+1}, we have to appeal to limits.)

We denote the parameter of a point by the corresponding small letter, e.g. $t(A_i)=a_i$. Since A_{i-1}, A_i and $A_{i+1}'=C_{i+1}^K$ are collinear points on Γ, we have (using (5·3)),

$$0\equiv a_{i-1}+a_i+a_{i+1}'\equiv a_{i-1}+a_i+c_{i+1}+\alpha \tag{6·2}$$

modulo the periods of the function f defining the parameter t. Since C_{i-1}', C_i, C_{i+1} are collinear points on Γ, and $C_{i-1}'^K=A_{i-1}$, we have

$$0\equiv c_{i-1}'+c_i+c_{i+1}\equiv a_{i-1}+c_i+c_{i+1}-\alpha. \tag{6·3}$$

On writing

$$B_i=A_i^K \tag{6·4}$$

we have $b_i\equiv a_i+\alpha$, so that from (6·2)

$$a_{i-1}+b_i+c_{i+1}\equiv 0. \tag{6·5}$$

so that, for each i, $A_{i-1}B_iC_{i+1}$ are collinear. A similar argument shows that $A_{i+1}B_iC_{i-1}$ are collinear. Also a comparison of (6·5) with (6·3) shows that

$$c_i = b_i + \alpha, \tag{6·6}$$

i.e.

$$C_i = B_i^K. \tag{6·7}$$

Hence from (6·5) we have

$$\begin{aligned} b_0 + b_1 + b_2 &\equiv (a_0 + \alpha) + b_1 + (c_2 - \alpha) \\ &\equiv a_0 + b_1 + c_2 \equiv 0, \end{aligned}$$

i.e. $B_0B_1B_2$ are collinear. Hence we have a Pascal configuration, with $A_0A_1A_2C_0C_1C_2$ lying on a conic and $B_0B_1B_2$ on a line.

In the particular case in which Λ is a harmonic homology H under which the cubic Γ is invariant, we know that the centre of the homology must be collinear with $C_iC_i^H = C_iC_i^\Lambda = C_iA_{-i}$ for all i. But $C_1A_{-1} = C_1A_2$ and $C_2A_{-2} = C_2A_1$ meet at B_0. Hence B_0 is the centre of H, and $A_0B_0C_0$ are also collinear.

Conversely, suppose that a set of points A_i, B_i, C_i form a Pascal configuration as above, and therefore is an associated set of points. We suppose that these are in general position, in the sense that no C_i lies on a side of the triangle $A_0A_1A_2$. Then we assert that there exists a linear transformation Λ and a general inversion G with fundamental triangle $A_0A_1A_2$ such that

$$B_i \text{ lies on } A_i^G, \quad C_i = B_i^{\Lambda G}, \quad A_i = C_{-i}^\Lambda, \tag{6·8}$$

and also that the A_i, B_i, C_i form a set of 9 associated points, and any cubic of the pencil through the A_i, B_i, C_i is invariant under $K = \Lambda G$.

This is easiest to prove analytically. Referred to the triangle of reference $A_0A_1A_2$ with suitable choice of co-ordinates the transformation G can be taken to be $(x_i) \mapsto (x_{-i}^{-1})$. The equation of the conic $\gamma = A_0A_1A_2C_0C_1C_2$ can be written as

$$\gamma: \sum_j (g_i x_i)^{-1} = 0, \tag{6·9}$$

for some non-zero g_i. The co-ordinates of C_0 can be written

$$C_0 = (g_0^{-1} : g_1^{-1}J_0^{-1} : g_2^{-1}j_0^{-1}) \tag{6·10}$$

for suitable choice of J_0, j_0 and (cyclically) similar expressions hold for C_1, C_2. Since each C_i lies on γ we have on substitution in (6·9)

$$1 + J_i + j_i = 0 \quad (i = 0, 1, 2). \tag{6·11}$$

The point B_0 is the intersection of A_1C_2 with A_2C_1 and hence is

$$B_0 = (g_0^{-1} : g_1^{-1}j_1 : g_2^{-1}J_2),$$

with (cyclically) similar expressions for B_1, B_2. If we now define the matrix

$$\mathbf{\Lambda} = \begin{bmatrix} g_0^2 J_2 j_1 & g_0 g_2 j_1 & g_0 g_1 J_2 \\ g_2 g_0 J_1 & g_2^2 J_1 j_0 & g_2 g_1 j_0 \\ g_1 g_0 j_2 & g_1 g_2 J_0 & g_1^2 J_1 j_0 \end{bmatrix} \tag{6·12}$$

and Λ, the corresponding linear mapping, then the truth of equations (6·8) is a matter of straightforward verification.

Furthermore, the set of 3 lines A_0C_2, A_1C_0, A_0C_1 together form a degenerate cubic with equation

$$\Gamma_{(1)}(\mathbf{x}) = {}_{\text{def}}\, \Pi(g_iJ_{i+1}x_i - g_{i+1}x_{i+1}) = 0,$$

and the set of lines A_0C_1, A_1C_2, A_2C_0 form a cubic

$$\Gamma_{(2)}(\mathbf{x}) = {}_{\text{def}}\, \Pi(g_ix_i - g_{i+1}J_ix_{i+1}) = 0.$$

The pencil of cubics defined by these two passes through all the points A_i, B_i, C_i and consists of all cubics of equations of the form

$$\Gamma_{(1)}(\mathbf{x}) + \mu\Gamma_{(2)}(\mathbf{x}) = 0. \tag{6·13}$$

It is easy to verify algebraically that each such cubic is invariant under the mapping $K = \Lambda G$.

The Mandel and Todd–Lyness transformations have the form HG, where H is a harmonic homology, and hence $HH = I$. We may therefore ask under what conditions does $\Lambda\Lambda = I$, i.e. $\mathbf{\Lambda}^2 = \mathbf{I}$ where $\mathbf{\Lambda}$ is given by (6·12). After a fair amount of manipulation this reduces to two conditions only,

$$g_0^2J_2 = g_1g_2J_0; \quad g_0^2j_1 = g_1g_2j_0, \tag{6·14}$$

which together imply that

$$J_2j_0 = J_0j_1, \tag{6·15}$$

i.e. $A_0B_0C_0$ are collinear.

Conversely, suppose that $A_0B_0C_0$ are collinear and (6·15) holds in some co-ordinate system $\boldsymbol{\xi}$ referred to the triangle $A_0A_1A_2$. Any other co-ordinate system $\mathbf{x}$ is given by

$$x_i = h_i\xi_i$$

for non-zero h_i. By choosing the h_i suitably we can ensure that (6·14) holds, which implies that $\mathbf{\Lambda}^2 = \mathbf{I}$, and Λ is a harmonic homology. It is then a matter of straightforward verification that the curves (6·13) are invariant under G, and hence also under $H = KG$.

If we have both $A_0B_0C_0$ and $A_1B_1C_1$ collinear, then we have

$$J_2j_0 = J_0j_1 = J_1j_2,$$

whence $A_2B_2C_2$ are also collinear. It then also follows from the same argument that the transformation K can be expressed in the form HG in 3 different ways, with each vertex A_i taken in turn as the centre of the corresponding generalized inversion G, and B_i as the centre of the corresponding homology H.

7. *An orbit of 5 points*

We now study in detail the parametrization of the cubic Γ_c^L given by (3·4), keeping c fixed. There will exist δ such that $t(P^L) = t(P) + \delta$. Also $t(P) + t(Q) + t(R) \equiv 0$ when P, Q, R are collinear.

 A quadratic mapping with invariant cubic curve. *Math. Proc. Camb. Phil. Soc.* **89** (1981), no. 1, 89–105. With C. A. B. Smith

Write $P_{-2}, P_{-1}, P_0, P_1, P_2$ for the five points $(1:0:-1)$, $(0:0:1)$, $(0:1:-1)$, $(0:1:0)$, and $(1:-1:0)$ respectively. These five points lie on Γ_c^L (which, for convenience, we refer to simply as Γ). Since $P_0^L = P_1$ and $P_2^L = P_{-2}$, we have (modulo the periods)

$$t(P_1) - t(P_0) \equiv \delta \equiv t(P_{-2}) - t(P_2). \tag{7·1}$$

The line $x_1 = 0$ meets Γ at P_{-1} and twice at P_{-2}, so that

$$t(P_{-1}) + 2t(P_{-2}) \equiv 0. \tag{7·2}$$

Similar sets of 3 collinear points on Γ are (P_1, P_2, P_2), (P_{-1}, P_{-1}, P_2), (P_1, P_1, P_{-2}), (P_{-1}, P_0, P_1) and (P_{-2}, P_0, P_2), giving rise to similar equations. The solution of these simultaneous equations is

$$t(P_r) \equiv r\delta, \quad \text{and} \quad 5\delta \equiv 0, \tag{7·3}$$

i.e. δ is a fifth part of a period. (Also it follows that $P_r^L = P_{r+1}$, suffixes modulo 5, so that we have an orbit of 5 points P_r.) We now produce explicit formulae for the parametrization. Change the co-ordinates from x to y, where

$$\left.\begin{aligned} x_0 &= y_1, \\ x_1 &= y_0 + \tfrac{1}{2}(c-1)\,y_1 + cy_2/\sqrt{8}, \\ x_2 &= y_0 + \tfrac{1}{2}(c-1)\,y_1 - cy_2/\sqrt{8}. \end{aligned}\right\} \tag{7·4}$$

The y-co-ordinates of the points P_r are

$$\left.\begin{aligned} P_{-2} &= (c^2: -2c: -\sqrt{8}), \\ P_{-1} &= (c:0: -\sqrt{8}), \\ P_0 &= (0:0:1), \\ P_1 &= (c:0:\sqrt{8}), \\ P_2 &= (c^2: -2c:\sqrt{8}). \end{aligned}\right\} \tag{7·5}$$

The equation of Γ becomes

$$\begin{aligned} c^2 y_2^2 y_0 &= 4c^2 y_1^3 + (2c^2 + 12c + 2)\,y_1^2 y_0 + (8c+8)\,y_1 y_0^2 + 8y_0^3 \\ &= 4c^2 (y_1 - \lambda_0 y_0)\,(y_1 - \lambda_1 y_0)\,(y_1 - \lambda_2 y_0), \end{aligned} \tag{7·6}$$

where the λ_r are the roots of the equation

$$4c^2\lambda^3 + (2c^2 + 12c + 2)\,\lambda^2 + (8c+8)\,\lambda + 8 = 0. \tag{7·7}$$

Following standard procedure (see, e.g., Neville (**11**, **12**)) set $\mu_r = \lambda_r - \lambda_0$ and

$$k = \sqrt{(\mu_2/\mu_1)}. \tag{7·8}$$

Then the parametrization is

$$\begin{aligned} y_0 &= 1, \\ y_1 &= \lambda_0 + \mu_1 \operatorname{ns}^2 t = \lambda_1 + \mu_1 \operatorname{cs}^2 t = \lambda_2 + \mu_1 \operatorname{ds}^2 t, \\ y_2 &= 2\mu_1^{\frac{3}{2}}.\operatorname{ns} t.\operatorname{cs} t.\operatorname{ds} t, \end{aligned} \tag{7·9}$$

where the elliptic functions ns, cs, ds have parameter k. Since these functions have simple poles at the origin, $t(P_0) = 0$, and hence $t(P_1) = t(P_0^L) = \delta$, so that

$$\lambda_0 + \mu_1 \operatorname{ns}^2\delta = \lambda_1 + \mu_1 \operatorname{cs}^2\delta = 0,$$

whence

$$\operatorname{cn}^2\delta = (\operatorname{cs}^2\delta)/(\operatorname{ns}^2\delta) = \lambda_1/\lambda_0. \tag{7·10}$$

Thus, given c, we can find $\operatorname{cn}^2\delta$ as the ratio of two roots of (7·7). But these relations can also conveniently be expressed in terms of a parameter θ. Set

$$\lambda_0 = -\theta^2/2. \tag{7·11}$$

Then (7·7) is a quadratic equation for c, of which one solution is

$$c = (2+\theta)^2/\theta^2(1+\theta) \tag{7·12}$$

(and the other solution is got by replacing θ by $-\theta$). But, given one root $\lambda_0 = -\frac{1}{2}\theta^2$ of (7·7), the other two can now be found straightforwardly. Write

$$\begin{aligned} p &= 2 + 2\theta + 2\theta^2 + \theta^3, \\ q &= 2(1+\theta)\sqrt{(1 - 2\theta^2 - \theta^3)}, \\ r &= \theta^2(2+\theta)^4/4(1+\theta), \\ k &= \sqrt{[(p-q-r)/(p+q-r)]}. \end{aligned}$$

Then the other two roots turn out to be

$$\lambda_1, \lambda_2 = r^{-1}(p \pm q)\,\lambda_0. \tag{7·13}$$

By use of (7·10) we then obtain

Theorem 7·1. If $\theta \neq 0$ or -2 then there exists δ such that, with p, q, r defined as above,

$$\begin{aligned} \operatorname{cn}^2(\delta; k) &= (p+q)/r, \\ \operatorname{dn}^2(\delta; k) &= (p-q)/r, \\ \operatorname{sn}^2(\delta; k) &= (r-p-q)/r, \end{aligned}$$

and 5δ is a period of these functions.

By considering the relation $t(P_2) = 2\delta$ we find similarly

$$\operatorname{sn}^2(2\delta; k) = (\theta+2)^2(r-p-q)/\theta^2 r = (\theta+1)\,c\,\operatorname{sn}^2(\delta; k). \tag{7·14}$$

The particular case $\theta = \tau^{-1} = (\sqrt{5}-1)/2$, where τ is the golden number, gives $c = \tau^5$, $p = \tau^3$, $q = 0$, $r = \tau^5/4$, $k = 1$. The elliptic functions then reduce to hyperbolic functions, so that

$$\operatorname{cn}^2\delta = \operatorname{sech}^2\delta = 4\tau^{-2}. \tag{7·15}$$

This corresponds to the fifth part of a period, $\delta = 2\pi i/5$, or the well-known relation for circular functions, $\cos 2\pi/5 = \frac{1}{2}\tau$. Similarly the equation (7·14) gives in this case $\operatorname{sn}^2 2\delta = 1 - 4\tau^2$, or $\operatorname{cn}^2 2\delta = \operatorname{sech}^2 2\delta = 4\tau^2$, in accordance with the relation

$$\cos 4\pi/5 = -1/2\tau.$$

The same results can alternatively be obtained by setting $\theta = -\tau$, which corresponds to $\delta = 4\pi i/5$.

8. *Acknowledgements*

We wish to thank cordially Professors H. S. M. Coxeter, P. Duval, J. G. Semple and W. K. Hayman for their interest and help, and both them and the referee for a number of suggestions which have improved the presentation and provided references to other relevant work.

REFERENCES

(1) Baker, H. F. *Principles of geometry*, vols 5, 6 (Cambridge University Press, 1933).
(2) Burnside, W. *Theory of groups of finite order*, 2nd ed. (Cambridge University Press, 1911).
(3) Coxeter, H. S. M. Affinely related polygons. *Abh. Math. Sem., Univ. Hamburg*, **34** (1969), 38–58.
(4) Coxeter, H. S. M. Frieze patterns. *Acta Arithmetica*, **28** (1971), 297–310.
(5) Duval, P. *Elliptic functions and elliptic curves*. London Mathematical Society Lecture Note Series, no. 9 (Cambridge University Press, 1973).
(6) Komissaruk, A. M. Foundations of affine geometry in the plane (in Russian). *Vyšejšaja škola*, Minsk, 1967.
(7) Lyness, R. C. Note 1581. *Math. Gaz.* **26** (1942), 62.
(8) Lyness, R. C. Note 1847. *Math. Gaz.* **29** (1945), 231.
(9) Mandel, S. P. H. The stability of a multiple allelic system. *Heredity*, **13** (1959), 289–302.
(10) Mulholland, H. P. and Smith, C. A. B. An inequality arising in genetical theory. *Amer. Math. Monthly* **66** (1959), 673–683.
(11) Neville, E. H. *Jacobian elliptic functions* (Oxford University Press, 1954).
(12) Neville, E. H. *Elliptic functions, a primer* (Pergamon Press, 1971).
(13) Scheuer, P. A. G. and Mandel, S. P. H. An inequality in population genetics. *Heredity* **13** (1959), 519–524.
(14) Segre, C. Le corrispondenze univoche sulle curve ellittiche. *Atti dell'Accademia delle Scienze, Torino* **24** (1889), 734–756.
(15) Semple, J. G. and Roth, L. *Introduction to algebraic geometry* (Oxford, Clarendon Press, 1949).
(16) Zeuthen, H. G. *Lehrbuch der Abzählenden Methoden der Geometrie* (Leipzig, Teubner, 1914).

127
TWISTORS FOR COSMOLOGICAL MODELS

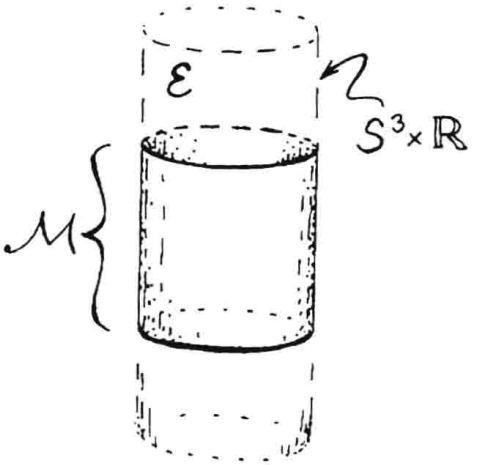
Twistor Newsletter **12** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

Twistors for cosmological models. *Twistor Newsletter* **12** (1981). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.3.2 Twistors for Cosmological Models *by R. Penrose* (TN 12, June 1981)

The Friedmann-Robertson-Walker [FRW] cosmological models are all conformally flat (as follows at once from their spatial isotropy, since the Weyl curvature, with its principal null directions, must vanish). Thus twistors (of the standard 'flat' variety) can be directly used to study their properties. These models differ from Minkowski space in that they have a different scale factor determining the metric, but also their global properties are different in the sense that they are conformal to different portions of the Einstein cylinder $\mathcal{E}$ ($\cong S^3 \times \mathbb{R}$). The space $\mathcal{E}$ is conformally the universal

covering space of $\mathbb{M}$ (compactified Minkowski space), while $\mathbb{M}$ itself may be regarded as a real quadric (signature $++---$) in $\mathbb{RP}^5$.



Let us first consider de Sitter space $\mathcal{M}$ and compare it with Minkowski space $\mathbb{M}^I$. The boundary of $\mathcal{M}$ (considered a conformal subset of $\mathcal{E}$) is a pair of spacelike hypersurfaces $\tilde{\mathscr{I}}$ and $\mathscr{I}$ which get identified with one another when $\mathcal{E}$ is wrapped back up to the quadric $\mathbb{M}$. This identified hypersurface turns out to be a hyperplane section of $\mathbb{M}$ in the projective space $\mathbb{RP}^5$. Removing this hypersurface from $\mathbb{M}$ we get a space conformal to $\mathcal{M}$. This is very similar to the case of $\mathbb{M}^I$, the only difference being that for $\mathbb{M}^I$, the hyperplane in $\mathbb{RP}^5$ touches $\mathbb{M}$, the contact point being i (or I). A hyperplane R in $\mathbb{RP}^5$ can be represented twistorially by a skew twistor $R_{\alpha\beta}$ which is 'real' in the sense that its twistor complex conjugate $\overline{R}^{\alpha\beta}$ and its dual $R^{\alpha\beta}$ are equal. We may also think of $R^{\alpha\beta}$ as representing a point in $\mathbb{RP}^5$, this point being the pole of R with respect to the quadric $\mathbb{M}$, but the point and hyperplane R represent equivalent information. The point lies on $\mathbb{M}$ iff its polar hyperplane R touches $\mathbb{M}$ iff the twistor $R_{\alpha\beta}$ is simple. Thus de Sitter space has an 'infinity twistor' $I^{\alpha\beta}$, with dual $I_{\alpha\beta}$, which is *real* in the above sense and which is *not simple*. In fact, for *de Sitter* space we always have

$$I_{\alpha\beta} I^{\alpha\beta} > 0,$$

while for *anti-de Sitter* space we should have

$$I_{\alpha\beta} I^{\alpha\beta} < 0,$$

the corresponding hypersurface $\mathscr{I}$ being now timelike. We can take the radius of the de Sitter pseudosphere to be

$$(I_{\alpha\beta} I^{\alpha\beta})^{-1/2}.$$

The de Sitter [anti-de Sitter] group is the subgroup of the twistor group $SU(2,2)$ which leaves $I^{\alpha\beta}$ invariant.

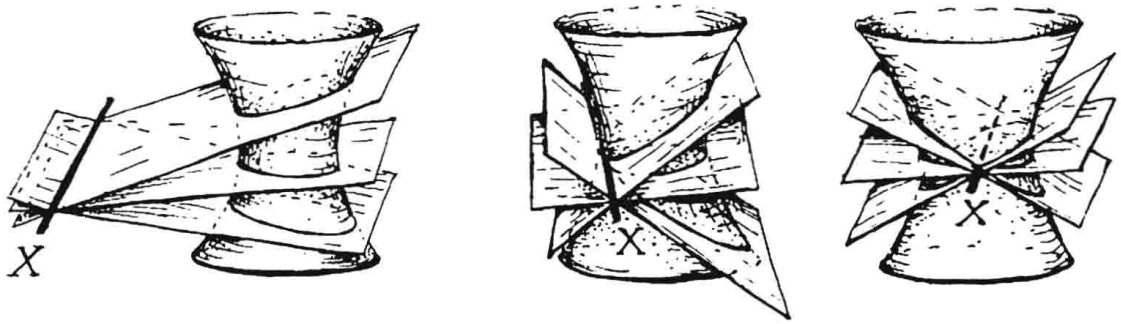
The particular FRW models just considered have the uncharacteristic property that they do *not* single out a unique 1-parameter family of hypersurfaces of homogeneity. In all other cases such a family *is* singled out, and they are determined in $\mathbb{RP}^5$ as the intersections of $\mathbb{M}$ with the hyperplanes of a pencil, i.e. by the family of skew twistors

$$\lambda I_{\alpha\beta} + \mu \tilde{I}_{\alpha\beta} \qquad (1)$$

Twistors for cosmological models. *Twistor Newsletter* **12** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

given by varying the ratio $\lambda : \mu$, where $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ are fixed non-proportional skew twistors. The members of the family (1) are to be all real (in the above twistor sense). One way of achieving this is to make $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ both real and to take $\lambda, \mu \in \mathbb{R}$. Another is to take $\tilde{I}_{\alpha\beta} = \bar{I}_{\alpha\beta}$ (= complex conjugate of dual of $I_{\alpha\beta}$) and $\lambda = \bar{\mu}$.

There are three allowed possibilities for the pencil (1): the axis X of the pencil (= the 3-plane common to all members of the pencil) may (i) miss $\mathbb{M}$ altogether, (ii) touch $\mathbb{M}$, or (iii) meet $\mathbb{M}$ in an S^2.



In case (i) the hypersurfaces of homogeneity are all S^3's ($k = 1$); in case (ii) they are all flat ($k = 0$); and in case (iii) the spacelike members of the family are all Lobachevsky (hyperbolic) spaces. There are also some disallowed cases in which the sections of $\mathbb{M}$ by the members of the pencil are all non-spacelike. In case (iii), there are two members of the pencil which touch $\mathbb{M}$ and we can, if desired, choose the corresponding members of the family (1) to be $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ themselves. In this case $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ will both be simple. They may or may not represent infinity for the FRW model, and they only could do so (for reasonable matter) if the cosmological constant vanishes. For non-vanishing cosmological constant we may prefer to choose $I_{\alpha\beta}$ (and/or $\tilde{I}_{\alpha\beta}$) actually to represent infinity. In that case we get $I_{\alpha\beta}I^{\alpha\beta}$ having the same sign as the cosmological constant (because $\mathscr{I}$ is spacelike for positive, null for zero, and timelike for negative cosmological constant). In case (i), we find that the members of the family which touch $\mathbb{M}$ would be conjugate complex, so if we choose these to be given by $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ we get $I_{\alpha\beta}$ complex and simple, and we can take $\tilde{I}_{\alpha\beta} = \bar{I}_{\alpha\beta}$. In case (ii) we find that only one member of the pencil touches $\mathbb{M}$ (giving (1) simple). In each one of these cases the symmetry group (leaving the homogeneity surfaces invariant) is the subgroup of $SU(2,2)$ leaving $I_{\alpha\beta}$ *and* $\tilde{I}_{\alpha\beta}$ invariant.

In models where there is a big bang (and/or big crunch) there will be a member of the pencil which corresponds to this. Thus we shall have a *bang twistor* $B_{\alpha\beta}$ (and/or a *crunch twistor* $C_{\alpha\beta}$) which is skew, twistor-real, and non-simple with

$$B_{\alpha\beta}B^{\alpha\beta} > 0 \quad (C_{\alpha\beta}C^{\alpha\beta} > 0)$$

since the bang (and/or crunch) is spacelike. If there is both a bang and a crunch there is the possibility that $B_{\alpha\beta}$ and $C_{\alpha\beta}$ turn out to be proportional—which would be the case with $k = +1$ if an observer 'just sees all the way around the universe' only at the moment that he encounters the crunch. (This is the case conformally identical with the de Sitter case, only here the boundary surfaces $\tilde{\mathscr{I}}$ and $\mathscr{I}$ at infinity are replaced by the singular boundaries of the bang and crunch, respectively). But in the general case, $B_{\alpha\beta}$ and $C_{\alpha\beta}$ would be non-proportional and could, if desired, be used in place of $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ in defining the pencil (1). An alternative—and this might be felt to be more appropriate in the case of a (say $k = -1$) model which starts from a bang and then expands to infinity—would be to use $B_{\alpha\beta}$ and $I_{\alpha\beta}$ (assuming they are non-porportional, as they would always be if $k = -1$).

Consider $k = \pm 1$ and make the choice (described earlier) where $I_{\alpha\beta}$ and $\tilde{I}_{\alpha\beta}$ are both simple ($\tilde{I}_{\alpha\beta} = \bar{I}_{\alpha\beta}$ if $k = +1$; $I_{\alpha\beta}$, $\tilde{I}_{\alpha\beta}$ both real if $k = -1$). We can scale things, assuming there is a bang, so that

$$B_{\alpha\beta} = \frac{1}{2}(I_{\alpha\beta} + \tilde{I}_{\alpha\beta}).$$

A point P of the space-time lies on the hypersurface given by (1) iff $(\lambda I_{\alpha\beta} + \mu\tilde{I}_{\alpha\beta})P^{\alpha\beta} = 0$ where $P^{\alpha\beta}$ is skew, simple and twistor-real, representing P. We can take this as

$$(\tau^{-1} I_{\alpha\beta} + \tau\tilde{I}_{\alpha\beta})P^{\alpha\beta} = 0$$

where $\tau = e^T$ for $k = -1$, and $\tau = e^{iT}$ for $k = +1$, with

$$T = \int_0 \frac{dt}{R(t)}.$$

Here t and $R = R(t)$ are the standard time and universe radius parameters for

$$ds^2 = dt^2 - R^2 d\Sigma^2$$

and $d\Sigma^2$ = squared metric of unit S^3 ($k = +1$) or unit Lobachevsky space ($k = -1$).

Editor's comment (1994). For further details see Penrose & Rindler (1986) *Spinors and space-time* Vol.2, CUP, §9.5 and Hurd, T.R. (1985) The projective geometry of simple cosmological models, Proc. Roy. Soc. **A397**, 233-43. See also §II.3.3.

128
PHYSICAL LEFT-RIGHT SYMMETRY AND GOOGLIES

Twistor Newsletter **12** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 2., 274–80. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

Introductory comments similar to those in relation to Chapter 125, are relevant here also.

Physical left-right symmetry and googlies. *Twistor Newsletter* **12** (1981). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. 274–80. 1995

§III.4.6 **Physical Left-Right Symmetry and Googlies** *by R. Penrose* (𝕋ℕ 12, July 1981)

In the framework of standard quantum mechanics there is, initially, nothing to choose between a position-space and a momentum-space representation. But clearly this does not imply that position and momentum are on an equal footing in physics. The Hamiltonians that one uses, for example, are generally without any symmetry with respect to

the interchange of position and momentum. Since pre-quantum physics had exhibited a manifest asymmetry between the concepts of position and momentum there was apparently no prior requirement for a formalism in which these two concepts would appear as completely interchangeable.

The situation with regard to twistor theory is rather the reverse of this. Here we have an analogous choice between a $\mathbb{T}$-representation and a $\mathbb{T}^*$-representation. But now interchanging $\mathbb{T}$ and $\mathbb{T}^*$ results in the interchange of left- and right-handedness (provided that the sign of i and the sense of time remain unaltered). And physics, for the most part *is* left-right symmetric!

On the other hand, it is part of the 'twistor philosophy' to regard the left-right asymmetries which are observed to be present in certain physical processes (P-violation in weak interactions; CP-violation in K^0-decay) as being manifestations of a preference that Nature might have for one or other of the spaces $\mathbb{T}$, $\mathbb{T}^*$. Accordingly, theory (and experiment) should ultimately reveal to us that it is one or the other of these spaces which is to be preferred as primary, while the other (like space-time itself) emerges as merely derived! In this way, the presence of left-right asymmetries in physics (somewhat analogously to the presence of position-momentum asymmetries in relation to the standard formalism) could be a manifestation of this primary/secondary status among the spaces $\mathbb{T}$, $\mathbb{T}^*$.

Our problem here, however, is rather the opposite one, namely, that most of known physics is actually left-right *symmetric*. A formalism based $\mathbb{T}$ on rather than on $\mathbb{T}^*$ (making, for the purpose of argument, a provisionally arbitrary choice) would seem to require an apparently miraculous built-in duality whereby almost every formula or logical argument would coexist also in a dual form, this dual version supplying the required twistor description of the space-reflection of the original version. Yet not everything in the theory should have an exact dual — in order that the space-asymmetric aspects of physics can be incorporated.

One of the simplest examples of a pair of formulae in which this kind of duality is playing a role occurs with the zero rest-mass field twistor contour integrals:

$$\phi_{A'\ldots L'} = \frac{1}{(2\pi i)^2}\oint \pi_{A'}\ldots\pi_{L'} f(ix^{QQ'}\pi_{Q'},\pi_{Q'})d^2\pi \qquad \text{(left-handed)} \qquad (1)$$

$$\phi_{A\ldots L} = \frac{\pm 1}{(2\pi i)^2}\oint \frac{\partial}{\partial\omega^A}\cdots\frac{\partial}{\partial\omega^L} f(ix^{QQ'}\pi_{Q'},\pi_{Q'})d^2\pi \quad \text{(right-handed)} \qquad (2)$$

The relation between (1) and (2) is a manifestation of the twistor quantization rule $\bar{Z}_\alpha \rightsquigarrow -\frac{\partial}{\partial Z^\alpha}$; i.e., $\bar{\omega}^{A'} \rightsquigarrow -\frac{\partial}{\partial \pi_{A'}}$, $\bar{\pi}_A \rightsquigarrow -\frac{\partial}{\partial \omega^A}$. Thus, in (2), instead of using the dual twistor $\bar{Z}_\alpha$, we use $\frac{\partial}{\partial Z_\alpha}$. In the particular cases of (1) and (2) in which the number of indices is equal to 4, we get space-time wave functions for the left-handed and right-handed parts of a linear graviton. For this, we could use the same f in (1) and in (2), where f has a part homogeneous of degree -6 and a part homogeneous of degree $+2$. Then f (or the cohomology class it determines) would be the $\mathbb{T}$-representation description of the wave function of the linear graviton.

Here 'linear graviton' refers to the standard linear spin-2 particle of the perturbation expansion of Einstein's nonlinear theory. According to the 'nonlinear graviton' philosophy, we ought really to be using the full Einstein theory to describe a single graviton, and not merely its linear approximation. In the case of a pure left-handed helicity state (helicity -2) we have at hand the so-called 'nonlinear graviton construction' in which the $+2$-homogeneous function f is regarded as providing an infinitesimal deformation of a portion ($\mathbb{T}^+$) of twistor space, but the twistor wave function is now more correctly a *finite* deformation. The philosophy requires that not only a purely left-handed state, but also the general mixed helicity state should be describable by an analogous such deformation, the googly case of pure right-handed helicity being a particular case. (It is perhaps unfortunate that the term 'nonlinear graviton' has become attached to the purely left-handed deformation of $\mathbb{T}^+$ — for understandable enough historical reasons. To be consistent with the cricketing terminology of the 'googly', I propose that the left-handed nonlinear graviton be referred to as the *'leg-break graviton'* when standard deformations of bits of $\mathbb{T}$ are involved! Then the term 'nonlinear graviton' could refer to a general mixed helicity state of positive frequency, satisfying the Einstein vacuum equations. It may be further remarked, incidentally, that although it would be very nice to have an ambitwistor-type description of general solutions of Einstein's equations, this would not be directly in line with the type of philosophy described here, where the solutions are thought of as providing nonlinear wave functions. An ambitwistor description would be more nearly analogous to a classical phase-space state, in which both position and momentum variables are required, rather than a quantum state which uses one *or* the other.

In order that a full twistor description, using $\mathbb{T}$ only, of Einstein's left-right symmetric theory be obtainable, we require that every essential concept involved in the standard

leg-break graviton should have a precise *dual* concept referring to the googly. In §III.4.3 a method of co-patching was indicated, according to which the googly information could be stored in a way dual to the normal patching together of a pair of portions of twistor space. This co-patching was achieved successfully only in the weak-field limit of the linear graviton — and even then, only with some imaginative leaps! What I am proposing here is a much more complete and rigorous programme of dualization, since only then can we really see how a left-right symmetric theory can emerge in the asymmetric setting of twistor space.

To illustrate what I have in mind, consider the exact sequence for flat twistor space $\mathbb{T}$:

$$0 \to \mathbb{S} \to \mathbb{T} \to \tilde{\mathbb{S}}^* \to 0$$

that features at the beginning of the googly graviton discussion. The idea is to consider not so much the spaces themselves, but the sheaves of holomorphic functions over these spaces. The projection $\mathbb{T} \to \tilde{\mathbb{S}}^*$ may be thought of in terms of the fact that holomorphic functions on $\tilde{\mathbb{S}}^*$ will lift back to functions on $\mathbb{T}$ which are annihilated by the operator $\frac{\partial}{\partial \omega^A}$. Indeed, we have the exact sequence of sheaves

$$0 \to \underset{\tilde{\mathbb{S}}^*}{\mathcal{O}} \to \mathcal{O} \xrightarrow{\partial/\partial\omega^A} \mathcal{O}_A \quad \left(\xrightarrow{\epsilon^{AB}\partial/\partial\omega^B} \mathcal{O} \to 0 \right)$$

(where $\mathcal{O}$ and $\mathcal{O}_A$ are sheaves over $\mathbb{T}$) defining $\underset{\tilde{\mathbb{S}}^*}{\mathcal{O}}$. Dually, we have

$$0 \leftarrow \underset{\mathbb{S}}{\mathcal{O}} \leftarrow \mathcal{O} \xleftarrow{\pi_{A'}} \mathcal{O}^{A'} \quad \left(\xleftarrow{\epsilon^{A'B'}\pi_{B'}} \mathcal{O} \leftarrow 0 \right)$$

defining $\underset{\mathbb{S}}{\mathcal{O}}$ and giving the sheaf version of the injection $\mathbb{S} \to \mathbb{T}$. This illustrates the general phenomenon that when we dualize, not only must we interchange Z^α and $\frac{\partial}{\partial Z^\alpha}$, but we must also reverse the directions of our sheaf maps.

For the googly graviton we require a deformed version of the injection $\mathbb{S} \to \mathbb{T}$ which does the dual of what, in the leg-break construction, preserves the Euler operator Υ and the 2-form $d\omega^A \wedge d\omega^B \epsilon_{AB}$ on each fibre of the projection $\mathcal{T} \to \tilde{\mathbb{S}}^*$. The dual of Υ is easy, since the exact sequence (for integer n, say)

$$0 \to \mathcal{O}(n) \to \mathcal{O} \xrightarrow{\Upsilon - n} \mathcal{O} \to \mathcal{O}(n) \to 0$$

dualizes to

$$0 \leftarrow \mathcal{O}(-n-4) \leftarrow \mathcal{O} \xleftarrow{-\Upsilon - n - 4} \mathcal{O} \leftarrow \mathcal{O}(-n-4) \leftarrow 0.$$

Thus the Euler operator is also well-defined (being replaced) merely by $-\Upsilon - 4$ on the googly twistor space.

The preservation of $d\omega^A \wedge d\omega^B \epsilon_{AB}$ is harder to dualize. One way is to use a co-patching which is dual to the generating function method of patching:

$$\omega^1 = \hat{\omega}^1 + \frac{\partial f}{\partial \omega^0}, \qquad \hat{\omega}^0 = \omega^0 + \frac{\partial f}{\partial \hat{\omega}^1},$$

where $f = f(\omega^0, \hat{\omega}^1, \pi_{0'}, \pi_{1'})$ is homogeneous of degree 2 and where $\hat{\pi}_{0'} = \pi_{0'}$, $\hat{\pi}_{1'} = \pi_{1'}$ (i.e., generating function $\omega^0\hat{\omega}^1 + f$). Writing out what is meant by such a patching in terms of a slightly confusing system of exact sequences and then dualizing as above, we apparently obtain the following concept of a function defined on the dual co-patched space: Consider functions $\Phi(\omega^A, \pi_{A'}, \hat{\pi}_{A'})$ (i.e., defined on 2-tailed twistor space) which satisfy the integro-differential relation

$$\frac{\partial \Phi}{\partial \pi_{A'}} - \frac{\partial \Phi}{\partial \hat{\pi}_{A'}} =$$
$$\frac{1}{(2\pi i)^2} \oint \oint \int_0^{\hat{\pi}_{1'}} \int_0^{\pi_{0'}} \Phi(\omega^A - \alpha^A, \pi_{0'} - \beta_{0'}, \pi_{1'}, \hat{\pi}_{0'}, \hat{\pi}_{1'} - \beta_{1'}) \beta^{A'} f(\alpha^A, \beta_{A'}) d^2\beta \wedge d^2\alpha .$$

Then we say that a pair of holomorphic functions $\{\Theta(\omega^A, \pi_{A'}), \Psi(\omega^A, \hat{\pi}_{A'})\}$ *reduced modulo* expressions of the form $\{\Phi(\omega^A, \pi_{A'}, 0), \Phi(\omega^A, 0, \hat{\pi}_{A'})\}$, where Φ satisfies (3), provides a definition of a function on the co-patched googly space.

All this seems rather obscure. It results from trying to make precise the reasoning put forward in §III.4.3 and applying it to the case of a *finite* co-deformation. As far as I can see, one ought to be able to use such expressions in conjunction with power series expansions about $\pi_{A'} = 0$ and $\hat{\pi}_{A'} = 0$ and the results checked term by term against the corresponding expansions (in ω^A and $\hat{\omega}^A$) of the leg-break construction. Assuming that all arguments have been correctly dualized, the googly expressions ought also to check out. (In a much simpler 1-dimensional model example, the 'co-Riemann sphere,' the analogous procedure does seem to work, though its geometric content is obscure. In this connection I should point out that the geometric meaning of a 'co-patching' in terms of a *'foliation'* on a product space — as had been suggested in §III.4.3 — is not really an accurate description in general. Instead, one seems to be presented with integro-differential equations, which are geometrically unpleasant!)

A more geometrical approach to the dualization of the 2-form $d\omega^A \wedge d\omega^B \epsilon_{AB}$ on each fibre of $\mathcal{T} \to \tilde{\mathbb{S}}^*$ can be given. Since $d\omega^A \wedge d\omega^B \epsilon_{AB}$ does not live naturally on $\mathcal{T}$

(the deformed $\mathbb{T}$ portion), but only on the fibres separately, it is more satisfactory to consider the bivector (degenerate Poisson structure)

$$\epsilon^{AB}\frac{\partial}{\partial\omega^A}\wedge\frac{\partial}{\partial\omega^B}=\underbrace{\partial_Z\wedge\partial_Z} \tag{4}$$

which does live on $\mathcal{T}$ and is the 'differential geometry dual' of $d\omega^A\wedge d\omega^B\epsilon_{AB}$ on each fibre separately. The operator (4) does not dualize directly (in the sense required here) because it does not operate on the sheaf $\mathcal{O}$, being something that acts on a pair of functions f, g on $\mathcal{T}$ to produce $\epsilon^{AB}\frac{\partial f}{\partial\omega^A}\frac{\partial g}{\partial\omega^B}=(\underbrace{\partial_Z f})(\underbrace{\partial_Z g})$. To express this suitably we must think of f and g as functions of different variables Z and $\tilde{Z}$ so that we can write the required structure $\underbrace{\partial_Z f(Z)}\partial_{\tilde{Z}}(\tilde{Z})=\underbrace{\partial_Z}\partial_{\tilde{Z}}\{f(Z)g(\tilde{Z})\}$ *restricted to the diagonal* $Z=\tilde{Z}$ on $\mathcal{T}\times\mathcal{T}$. The fact that $f(Z)g(\tilde{Z})$ is a product plays no role here. (The completed linear span of such products provides the general holomorphic function on $\mathcal{T}\times\mathcal{T}$.) Our structure can be thought of as a map taking functions of two variables to functions of one variable ('functions' being always taken in the local sheaf sense): $F(\ ,\ )\mapsto f(\ )$, where $f(Z)=\underbrace{\partial_Z}\partial_{\tilde{Z}}F(Z,\tilde{Z})|_{\tilde{Z}=Z}$.

Writing $\mathcal{O}$, now, for the sheaf of holomorphic functions on $\mathcal{T}\times\mathcal{T}$ (or $\mathbb{T}\times\mathbb{T}$), we can express our structure as the map '$- - \twoheadrightarrow$' in the diagram (with exact row and commuting triangle):

$$\begin{array}{ccccccccc} & & & & \mathcal{O} & & & & \\ & & & & \downarrow{\scriptstyle \underbrace{\partial_Z}\partial_{\tilde{Z}}} & \searrow & & & \\ \mathcal{O}. & \xrightarrow[Z\cdot-\tilde{Z}\cdot]{} & \mathcal{O} & \longrightarrow & \underset{\tilde{Z}=Z}{\mathcal{O}} & \to & 0\,. & & \end{array}$$

Dualizing, we get:

$$\begin{array}{ccccccccc} & & & & \mathcal{O} & & & & \\ & & & & \uparrow{\scriptstyle \overrightarrow{Z}\overrightarrow{\tilde{Z}}} & \nwarrow & & & \\ \mathcal{O}. & \xleftarrow[-\partial_Z+\partial_{\tilde{Z}}]{} & \mathcal{O} & \longleftarrow & \underset{Z+\tilde{Z}}{\mathcal{O}} & \leftarrow & 0\,. & & \end{array}$$

Our dual structure therefore takes the form of a map from functions of one variable to functions of two variables: $f(\ )\mapsto F(\ ,\ )$, where $F(Z,\tilde{Z})=\overrightarrow{Z}\overrightarrow{\tilde{Z}}f(Z+\tilde{Z})$. This seems a little odd, at first, since '$f(Z+\tilde{Z})$' seems to imply a linear structure which would be already stronger than the structure we seek. However, this is not so because the 'product space' over which $\mathcal{O}$ is now defined does not allow a canonical concept of 'holding Z constant' or of 'holding $\tilde{Z}$ constant'. All we are allowed to do is to 'put $Z=0$' or 'put $\tilde{Z}=0$' (so the dual product is more like a disjoint union — which is 'categorically' more accurate). However, putting $Z=0$ or $\tilde{Z}=0$ in $\overrightarrow{Z}\overrightarrow{\tilde{Z}}f(Z+\tilde{Z})$ simply yields zero.

 Physical left-right symmetry and googlies. *Twistor Newsletter* **12** (1981). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. 274–80. 1995

Differentiating first with respect to Z or $\tilde{Z}$, we get *two* maps like $f(Z) \mapsto \tilde{Z} \rceil f(Z)$. Though more work needs to be done to understand this, it seems that this map refers to a 'blow up' or 'resolution' of the line I, whereas the existence of *two* such maps seems to imply that I occurs in two copies, so we get a non-Hausdorff projective space. Both these properties are desirable in this context.

129
COHOMOLOGY AND MASSLESS FIELDS

Commun. Math. Phys. **78** (1980/81), no. 3, 305–51.
With M. G. Eastwood and R. O. Wells, Jr.

This is the serious mathematical write-up of the twistor-cohomological representation of massless fields, following up on the primitive ideas first introduced in Chapter 89, and carrying the correspondence much further as an abstract relationship that encompasses many different specific representations.

Cohomology and Massless Fields*

Michael G. Eastwood[1], Roger Penrose[1], and R. O. Wells[2], Jr.

1 Mathematical Institute, Oxford University, Oxford OX1 3LB, England, and
2 Department of Mathematics, Rice University, Houston, TX 77001, USA

Introduction

The geometry of twistors was first introduced in Penrose [28]. Since that time it has played a significant role in solutions of various problems in mathemetical physics of both a linear and nonlinear nature (cf. Penrose [29], Penrose [35], Ward [48], and Atiyah–Hitchin–Drinfeld–Manin [2], see Wells [52] for a recent survey of the topic with a more extensive bibliography). The major role it has played has been in setting up a general correspondence which translates certain important physical field equations in space-time into holomorphic structures on a related complex manifold known as *twistor space*. The purpose of this paper is to give a rigorous discussion of this correspondence for the case of the linear massless free-field equations, including Maxwell's source-free equations, the wave equation, the Dirac–Weyl neutrino equations, and the linearized (weak-field limit of) Einstein's vacuum equations. These equations may also be analyzed from this point of view on a background provided by the nonlinear Yang–Mills or Einstein equations in the (anti-) self-dual case. The correspondence is effected by an integral-geometric transform, which transforms complex-analytic data on twistor space to solutions of the linear massless field equations, and is, in fact, a generalization of the classical Radon transform, which is discussed further below.

The motivation for finding such a correspondence in general is that it forms an essential part of the "twistor programme" according to which one attempts to eliminate the equations of physics by deriving them from the rigidity of complex geometry and holomorphic functions (see, e.g. Penrose [38]). It is, in fact, rather remarkable the extent to which it is possible to achieve this. Success apparently comes about because in twistor-space descriptions the information is "stored" nonlocally. The (local) value of a field at a point in space-time depends upon the way that the holomorphic structure in the twistor-space is fitted together in the large. So sheaf cohomology and function theory of several complex variables turn out to be the appropriate tools in the twistor framework. It is hoped that, as part of the general twistor programme, some deeper insights may eventually be gained as to the interrelation between quantum mechanics or quantum field

* Research supported by the National Science Foundation, the Institute for Advanced Study, and the Vaughn Foundation

theory (which depend crucially upon complex function theory) and classical space-time structure.

Our discussion for the most part will be restricted to the case of analytic linear massless fields on flat space-time and will be carried out in terms of the compactified complexification $\mathbb{M}$ (often denoted $\mathbb{CM}$ elsewhere) of Minkowski-space. For this purpose, $\mathbb{M}$ will simply be *defined* as the Grassmannian of linear 2-spaces in twistor space, which is a complex vector space $\mathbb{T}$ of dimension 4. We obtain a number of theorems relating analytic cohomology of portions of $\mathbb{T}$, or rather of the corresponding projective 3-space $\mathbb{P}$, to the solutions of the relevant field equations in the corresponding regions of $\mathbb{M}$.

In the case of such massless fields the above-mentioned correspondence was first given in terms of a contour integral expression which yields general analytic solutions of the field equations when a holomorphic function, defined on a suitable domain, is inserted into the integrand (for right-handed fields see Penrose [29], [30], [33], and Penrose–MacCallum [32]; for left-handed fields see Penrose [34] or Hughston [23]). (For the wave-equation a very closely related integral expression had been found much earlier by Bateman [4].) Only comparatively recently (Penrose [36]) was it realized that these holomorphic functions should be interpreted as Čech cocycles representing elements of cohomology groups.[1] A definition of the transform (for right-handed fields) using the Dolbeault representation of analytic cohomology is described in Woodhouse [55] and Wells [52]. In Hughston [21] and Ward [49] it was shown how elements of cohomology in the left-handed case give rise to potentials for fields and a method of producing the field itself more directly related to the integral formulae is described in Penrose [37] and more invariantly in Wells [53]. The question of whether the transform is bijective was considered in Penrose [37], Lerner [24], Eastwood [8], and Wells [52]. It was shown recently that the transform, in a special case, is also a special case of a generalization of the Radon transform due to Gindikhin and Henkin (see [15]). Integral formulae for background coupled fields have been given in Ward [47]. In this paper we make no explicit reference to the earlier contour integral descriptions but give our discussion entirely in terms of the mathematically more satisfactory cohomological language.

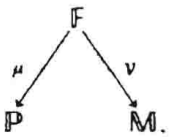
One example of massless field equations is the linearized Einstein equations (the case of helicity ± 2). By the theory developed in this paper the cohomology group $H^1(\mathbb{P}^+, \mathcal{O}(2))$ represents holomorphic solutions of helicity -2 of these equations on the forward tube $\mathbb{M}^+$ in complexified Minkowski space. This cohomology group is precisely the infinitesimal parameter space for the deformations of structure of the fibration $\mathbb{P}^+ \to \mathbb{P}_1$ which appear in the nonlinear graviton construction (Penrose [35]), and which give the general (local) holomorphic anti-self-dual solution of Einstein's equations. These solutions to the massless field equations in terms of cohomology groups relate in a similar manner to solutions of other nonlinear problems and, it is hoped, will give insight into as yet unsolved nonlinear problems (e.g., the self-dual or general solution of

1 The second author would like to express his gratitude to M. F. Atiyah for suggestions ultimately leading to this interpretation

Einstein's equations). In addition, there are algorithms for generating solutions of nonlinear field equations with specific cohomological solutions of associated linear equations as a principal part of the algorithm (cf. Atiyah–Ward [3], Ward [50]).

In the recent solution of the self-dual Yang–Mills equations on S^4 a key ingredient was the representation of specific cohomology groups on $\mathbb{P}_3$ in terms of solutions of conformally invariant elliptic differential equations on S^4 (see Atiyah–Hitchin–Drinfeld–Manin [2], Drinfeld–Manin [7], Hitchin [19], and Rawnsley [42]). This is a special case of our general method of representing solutions of linear field equations in terms of cohomology as discussed in Sect. 9.

We will now give a brief outline of the paper. In Sect. 1 we set up the basic twistor correspondence between projective twistor space (denoted by $\mathbb{P}$) and the natural complexification of compactified Minkowski space (denoted by $\mathbb{M}$). In particular we set up an important relative de Rham sequence on the flag manifold $\mathbb{F}$ which gives the correspondence between $\mathbb{P}$ and $\mathbb{M}$ via a double fibration of the following type



The basic integral transform of data on $\mathbb{P}$ to data on $\mathbb{M}$ is achieved by pulling back cohomology classes from (open subsets of) $\mathbb{P}$ to $\mathbb{F}$ by the mapping μ, and then "integrating this pullback class over the fibers" appropriately, obtaining (vector-valued) functions on (open subsets of) $\mathbb{M}$ which will satisfy certain field equations. We do this in several steps. First, we study the pullbacks of local data (inverse image sheaves and the relative deRham sequence in Sect. 1). Then we study the problem of "integration over the fibre" for the mapping ν. This involves direct image sheaves and a fundamental spectral sequence for a fibration due to Leray, but is a direct translation of the more classical notion of fibre-wise integration of a differential form into an algebraic and coordinate-free language which makes the group-invariance properties of this process manifest. The direct images of the pullbacks of the various powers of the hyperplane-section bundle on $\mathbb{P}$ are identified with *spinor sheaves* on $\mathbb{M}$, sheaves whose sections, locally, are classical spinor fields of various kinds. We also identify in some detail the invariant differential operators coming from the relative deRham complex with specific classical differential operators acting on spinor fields on $\mathbb{M}$. It is this identification which allows us to make the transition between complex geometry and mathematical physics. One consequence is that the massless field equations become transparently conformally invariant. In Sect. 3 we study the global problem of preserving information when pulling back cohomology classes from $\mathbb{P}$ to $\mathbb{F}$ by the mapping μ. This turns out to be a purely topological problem involving the fibres of the mapping $\mu : \hat{W} \to W$, where $W \subset \mathbb{P}$ and $\hat{W} \subset \mathbb{F}$ are open sets to which the mapping μ is restricted. It turns out that the pullback sheaves and the sheaves to which the direct image theory applies don't coincide, but they are related by the relative deRham exact sequences, and the corresponding cohomology

groups are related by a standard spectral sequence. This is discussed in Sect. 4. This completes the set of tools necessary to derive the massless field equations and their solutions. In Sect. 5 we obtain the isomorphism, for $s > 0$, $\mathscr{P}: H^1(\mathbb{P}^+, \mathcal{O}_{\mathbb{P}}(-2s-2)) \xrightarrow{\cong} \{$holomorphic right-handed massless fields of helicity s on $\mathbb{M}^+\}$. This isomorphism is equivariant with respect to the action of $U(2,2)$ on $\mathbb{T}$. We obtain this also for more general geometric settings. In Sect. 6 we derive a similar result for massless fields of helicity zero, where the second-order wave equation comes in. The spectral sequence theory shows why first-order equations arise for non-zero helicity, while second-order equations arise for helicity zero. In fact, all of the differential equations which are solved are first derived in an invariant form from the twistor geometry (and the spectral sequence on $\mathbb{F}$), then they are compared with the classical (noninvariant) forms, and then the solutions are produced by the integral transform. In Sect. 7 the left-handed solutions are described in terms of potentials which are transforms of certain analytic data. This is compared in Sect. 8 with a direct approach (power series expansions of cohomology classes about a given complex line in projective twistor space) given in Wells [53]. In Sect. 9 we generalize our results to background coupled fields, where the background potential is described in twistor space by means of a holomorphic line bundle (or more generally a holomorphic vector bundle) as was originated by Ward ([46], [47], [48]). Finally, in Sect. 10 we discuss various ramifications and possible extensions of the methods and ideas developed in this paper.

1. The Twistor Correspondence and a Relative de Rham Complex

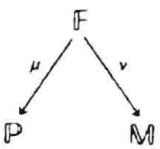
Let $\mathbb{T}$ be the space of *twistors*. $\mathbb{T}$ is, by definition, a complex vector-space of dimension 4 with an Hermitian form Φ of signature $++--$. We define

$$\mathbb{F}_{d_1,\ldots,d_r} := \{(L_1, \ldots, L_r) : L_1 \subset \ldots \subset L_r \text{ is a sequence of linear subspaces of } \mathbb{T} \text{ with } \dim_{\mathbb{C}} L_j = d_j.\}$$

and set

$$\mathbb{F} := \mathbb{F}_{1,2}, \quad \mathbb{P} := \mathbb{F}_1, \quad \mathbb{M} := \mathbb{F}_2.$$

These three compact complex manifolds (called *flag manifolds*, cf. Wells [51]) are naturally linked by the double fibration



(1.1)

where $\mu(L_1, L_2) = L_1$ and $\nu(L_1, L_2) = L_2$. The *twistor correspondence*[2] is the

2 If $\mathbb{M}$ is embedded in $\mathbb{P}_5$ by the Plücker coordinates then this is often referred to as the *Klein correspondence*

set-theoretic assignment (cf. Wells [52])

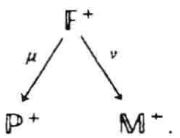
$$Z \in \mathbb{P} \longmapsto \hat{Z} := \nu(\mu^{-1}(Z)) \subset \mathbb{M}$$

$$z \in \mathbb{M} \longmapsto L_z := \mu(\nu^{-1}(z)) \subset \mathbb{P}$$

We note that $\mathbb{P} \simeq \mathbb{P}_3(\mathbb{C})$, three-dimensional projective space, $\mathbb{M} \simeq G_{2,4}(\mathbb{C})$, the Grassmannian manifold of 2-planes in 4-space, $L_z \simeq \mathbb{P}_1(\mathbb{C})$, a complex projective line embedded in $\mathbb{P}$, and $\hat{Z} \simeq \mathbb{P}_2(\mathbb{C})$, and is referred to as the *α-plane* in $\mathbb{M}$ determined by Z (the *β-planes* in $\mathbb{M}$ are determined by the projective hyperplanes in $\mathbb{P}$, but won't play a role in this paper). With a slight change in notation (1.1) is the basic double fibration used in [52] and we refer the reader to this article for more details and background references concerning this correspondence. Those already familiar with the basics of twistor theory will recognize $\mathbb{P}$ as *projective twistor-space*, $\mathbb{M}$ as *complexified compactified Minkowski-space*, and the correspondence space $\mathbb{F}$ as the (*dual*) *projective primed spin-bundle* over $\mathbb{M}$ (cf. Penrose MacCallum [32]).

Our object in this section is to describe certain natural sheaves on $\mathbb{P}$ and $\mathbb{F}$ and how they are linked. These sheaves will in some sense provide the basic local data for the integral-geometric transform developed in the later sections.

The quadratic form Φ determines distinguished open subsets of $\mathbb{P}, \mathbb{F}$, and $\mathbb{M}$, to be denoted by $\mathbb{P}^+$, $\mathbb{M}^-$, etc. just as in [52]. In particular the analysis in [52] concentrated on

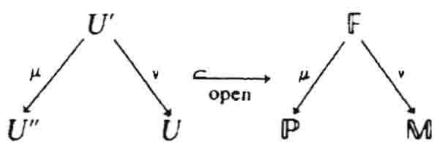


We shall refer to these subsets explicitly when we need them but we want to consider more general open sets related in the same manner but not necessarily SU(2, 2)-invariant. So let U be an arbitrary open subset of $\mathbb{M}$ and define

$$U' := \nu^{-1}(U)$$

$$U'' := \mu(U') \quad \text{(also occasionally denoted by } L_U\text{)}.$$

The analytical and topological properties of these sets will determine the precise nature of the transform to be developed later. We have the open inclusion of the two diagrams



(1.2)

with μ and ν surjective mappings of maximal rank.

We now want to introduce on $\mathbb{F}$ a complex of differential forms along the fibres of μ. First we introduce some standard notation.[3] If X is a complex manifold

3 See, e.g. Wells [51]. We refer to this reference for standard concepts concerning complex manifolds used in this paper

we let $\mathcal{O}_X$ be the sheaf of holomorphic functions on X and let Ω^p_X be the sheaf of holomorphic p-forms on $X, p \geqq 0$, noting that $\Omega^0_X = \mathcal{O}_X$. Then we have

$$0 \longrightarrow \mathbb{C} \longrightarrow \mathcal{O}_X \xrightarrow{d} \Omega^1_X \xrightarrow{d} \Omega^2_X \longrightarrow \ldots,$$

the exact deRham sequence of sheaves where d is the usual exterior differentiation operator. Define the sheaf Ω^1_μ of *relative 1-forms* with respect to the fibration μ by the exact sequence

$$\mu^*\Omega^1_{\mathbb{P}} \longrightarrow \Omega^1_{\mathbb{F}} \xrightarrow{\pi_\mu} \Omega^1_\mu \longrightarrow 0$$

i.e., Ω^1_μ is the quotient sheaf and π_μ is the usual quotient mapping. Then we set $d_\mu = \pi_\mu \circ d$. We shall write out explicit representations for the sheaf Ω^1_μ and the differential operator d_μ shortly, but first we extend this concept to higher order forms and to vector-bundle-valued differential forms. Let Ω^2_μ be the sheaf of *relative 2-forms* defined by the exact sequence

$$\mu^*\Omega^1_{\mathbb{P}} \wedge \Omega^1_{\mathbb{F}} \longrightarrow \Omega^2_{\mathbb{F}} \longrightarrow \Omega^2_\mu \longrightarrow 0$$

and let

$$d_\mu : \Omega^1_\mu \longrightarrow \Omega^2_\mu$$

be the induced exterior derivative. Let $\mu^{-1}\mathcal{O}_{\mathbb{P}}$ denote the topological inverse image of the sheaf $\mathcal{O}_{\mathbb{P}}$ i.e. the sections of $\mu^{-1}\mathcal{O}_{\mathbb{P}}$ are locally simply pullbacks by μ of sections of $\mathcal{O}_{\mathbb{P}}$ or, in other words, $\mu^{-1}\mathcal{O}_{\mathbb{P}}$ is the subsheaf of $\mathcal{O}_{\mathbb{F}}$ consisting of those functions locally constant on the fibres of μ.

Lemma 1.1. *The sequence*

$$0 \longrightarrow \mu^{-1}\mathcal{O}_{\mathbb{P}} \longrightarrow \mathcal{O}_{\mathbb{F}} \xrightarrow{d_\mu} \Omega^1_\mu \xrightarrow{d_\mu} \Omega^2_\mu \longrightarrow 0 \tag{1.3}$$

is exact.

Before we prove this lemma we shall give an easy generalization. Suppose that V is a holomorphic vector-bundle over U, an open subset of $\mathbb{P}$, and let μ^*V be the pullback bundle over $\mu^{-1}(U)$. The bundle μ^*V may be described by transition functions which are constant along the fibres of μ. We can tensor (1.3) with $\mu^{-1}\mathcal{O}_{\mathbb{P}}(V)$, the sheaf of germs of sections of μ^*V which are constant along the fibres of μ, to obtain (since $\mathcal{O}_{\mathbb{F}} \otimes_{\mu^{-1}\mathcal{O}_{\mathbb{P}}} \mu^{-1}\mathcal{O}_{\mathbb{P}}(V)$ is canonically isomorphic to $\mathcal{O}_{\mathbb{F}}(\mu^*V)$)

$$\begin{aligned} 0 \longrightarrow \mu^{-1}\mathcal{O}_{\mathbb{P}}(V) \longrightarrow \mathcal{O}_{\mathbb{F}}(\mu^*V) &\xrightarrow{d_\mu} \mathcal{O}_{\mathbb{F}}(\mu^*V) \otimes_{\mathcal{O}_{\mathbb{F}}} \Omega^1_\mu \\ &\xrightarrow{d_\mu} \mathcal{O}_{\mathbb{F}}(\mu^*V) \otimes_{\mathcal{O}_{\mathbb{F}}} \Omega^2_\mu \longrightarrow 0. \end{aligned} \tag{1.4}$$

Note that the differential operators d_μ in (1.4) make sense since they annihilate the transition functions of μ^*V. For convenience of notation we shall write

$$\Omega^0_\mu := \mathcal{O}_{\mathbb{F}}$$
$$\Omega^p_\mu(V) := \mathcal{O}_{\mathbb{F}}(\mu^*V) \otimes_{\mathcal{O}_{\mathbb{F}}} \Omega^p_\mu$$

and we can rewrite (1.4) as

$$0 \longrightarrow \mu^{-1}\mathcal{O}_{\mathbb{P}}(V) \longrightarrow \Omega^0_\mu(V) \xrightarrow{d_\mu} \Omega^1_\mu(V) \xrightarrow{d_\mu} \Omega^2_\mu(V) \longrightarrow 0 \tag{1.5}$$

For the special case of $V = H^n$, the n-th power of the hyperplane section bundle (cf. Wells [52]) we will write $\Omega^p_\mu(n)$ instead of $\Omega^p_\mu(H^n)$. This is consistent with the usual notation for "twisting" bundles on projective space. Thus we obtain in this case, for $n \in \mathbb{Z}$,

$$0 \longrightarrow \mu^{-1}\mathcal{O}_{\mathbb{P}}(n) \longrightarrow \Omega^0_\mu(n) \xrightarrow{d_n} \Omega^1_\mu(n) \xrightarrow{d_n} \Omega^2_\mu(n) \longrightarrow 0 \qquad (1.6)$$

This exact sequence of sheaves contains the basic local data for representing solutions of the massless field equations on $\mathbb{M}$. We note that (1.6) remains invariant under the action of $GL(\mathbb{T})$ on $\mathbb{T}$ and its induced action on $\mathbb{P}$, $\mathbb{F}$, and $\mathbb{M}$ as well as on H^n. We shall need to "globalize" this data using cohomology and then "push it down to $\mathbb{M}$" using direct images (cf. Sect. 2).

Now we shall proceed towards the proof of Lemma 1.1 and give explicit representations for these sheaves and differential operators. Actually the statement of Lemma 1.1 is local at a point of $\mathbb{F}$ and the proof will depend only on the fact that μ has maximal rank so that, by the implicit function theorem, one can always choose local coordinates in which μ is simply a coordinate projection (i.e. part of $\mathbb{C}^5 = \mathbb{C}^3 \times \mathbb{C}^2 \to \mathbb{C}^3$). However, it is useful at this stage to introduce specific coordinates tied to the spinor structure of the twistor-space $\mathbb{T}$. The coordinates have important physical interpretations (cf. Penrose–MacCallum [32]) and moreover the maximal rank condition will be obvious. Choose coordinates (Z^0, Z^1, Z^2, Z^3) for $\mathbb{T}$ and set

$$(Z^0, Z^1) = (\omega^0, \omega^1)$$
$$(Z^2, Z^3) = (\pi_{0'}, \pi_{1'})$$

so that $(\omega^A, \pi_{A'})$ are the components of a pair of *spinors* representing the *twistor* Z^α (cf. [32] and [52]). Indeed, we shall often write $\mathbb{T}^\alpha$ instead of $\mathbb{T}$ and $\mathbb{T}_\alpha$ for the dual of $\mathbb{T}^\alpha$. These coordinates are chosen so that

$$\Phi(Z^\alpha) = \omega^A \bar{\pi}_A + \bar{\omega}^{A'} \pi_{A'} \quad \text{(summation convention)}$$

or, equivalently, in matrix form

$$\Phi = \begin{bmatrix} 0 & I_2 \\ I_2 & 0 \end{bmatrix}.$$

We choose *affine coordinates* z^a for $\mathbb{M}^I$, the *affine part* of $\mathbb{M}$, where $\mathbb{M}^I$ is defined by the open inclusion

$$\begin{aligned} \mathbb{C}^4 \cong \mathbb{C}^{2\times 2} &\xrightarrow{\simeq} G^I_{2,4}(\mathbb{C}) \simeq \mathbb{M}^I \subset \mathbb{M} \\ z^a \leftrightarrow z^{AA'} &\overset{\omega}{\longmapsto} \operatorname{span} \begin{bmatrix} iz^{AA'} \\ I_2 \end{bmatrix}, \end{aligned} \qquad (1.7)$$

under the biholomorphism $G_{2,4}(\mathbb{C}) \simeq \mathbb{M}$. We see that $\mathbb{M}^I$ is defined by the set of 2-dimensional subspaces of $\mathbb{T}^\alpha$ which are linear graphs over the $\{\pi_{A'}\}$-coordinate plane, i.e. planes in $\mathbb{T}^\alpha$ of the form $\omega^A = iz^{AA'}\pi_{A'}$. $\mathbb{M}$ is covered by a finite number of such affine coordinate systems (in fact, five are needed). We note that $\mathbb{M}^+$ is the subset of $\mathbb{M}^I$ where, writing $z^a = x^a - iy^a$, y^a is timelike and future-pointing, and that *real (affine) Minkowski-space* is given when $y^a = 0$ (or equivalently,

when $z^{AA'} = x^{AA'}$ is Hermitian) (cf. [52]). We let $\mathbb{F}^I := \nu^{-1}(\mathbb{M}^I)$ and $\mathbb{P}^I := \mu(\mathbb{F}^I)$ be the associated "*affine*" portions of $\mathbb{F}$ and $\mathbb{P}$ respectively (although they are not affine in the sense of algebraic geometry). It is easy to describe the naturally associated coordinates on $\mathbb{P}^I$ and $\mathbb{F}^I$. By varying our choice of affine coordinates for $\mathbb{M}$ we could obtain in this way a complete system of linked coordinate charts on the three manifolds. First we observe that

$$\mathbb{F}^I \simeq \mathbb{M}^I \times \mathbb{P}_1 \text{ with coordinates } (z^{AA'}, [\pi_{A'}])$$
$$\mathbb{P}^I = \mathbb{P} - I,$$

where I is the projective line in $\mathbb{P}$ given by[4]

$$I = \{[\omega^A, \pi_{A'}] \in \mathbb{P} : \pi_{A'} = 0\}$$

and, as usual, [ ] denotes homogeneous coordinates on a projective space. Moreover,

$$\mathbb{F}^I = \mathbb{F}^I_0 \cup \mathbb{F}^I_1$$
$$\mathbb{P}^I = \mathbb{P}^I_0 \cup \mathbb{P}^I_1$$

where

$$\mathbb{P}^I_0 := \{[\omega^A, \pi_{A'}] \in \mathbb{P} : \pi_{0'} \neq 0\}$$
$$\mathbb{P}^I_1 := \{[\omega^A, \pi_{A'}] \in \mathbb{P} : \pi_{1'} \neq 0\}$$

and

$$\mathbb{F}^I_0 := \mu^{-1}(\mathbb{P}^I_0) = \mathbb{M}^I \times (\mathbb{P}_1 - \{\pi_{0'} = 0\})$$
$$\mathbb{F}^I_1 := \mu^{-1}(\mathbb{P}^I_1) = \mathbb{M}^I \times (\mathbb{P}_1 - \{\pi_{1'} = 0\}).$$

On $\mathbb{P}^I_0$ we introduce affine coordinates (q^0, q^1, r) by

$$q^A := \omega^A/\pi_{0'}, \quad r := \pi_{1'}/\pi_{0'}$$

and on $\mathbb{F}^I_0$ the coordinates

$$q^A := (iz^{AA'}\pi_{A'})/\pi_{0'}, \quad r := \pi_{1'}/\pi_{0'}$$
$$s^A := -z^A_{0'} = z^{A1'}.$$

With these coordinates $\mu|_{\mathbb{F}^I_0}$ is reduced to a projection

$$\begin{array}{ccccc} \mathbb{F}^I_0 \simeq & \mathbb{C}^3 \times \mathbb{C}^2 & \ni & (q^A, r, s^A) \\ \downarrow & \downarrow & & \downarrow \\ \mathbb{P}^I_0 \simeq & \mathbb{C}^3 & \ni & (q^A, r) \end{array}$$

so that s^A may be regarded as coordinates along the fibres of μ and d_μ may be realized explicitly as

$$d_\mu f = \frac{\partial f}{\partial s^0} d_\mu s^0 + \frac{\partial f}{\partial s^1} d_\mu s^1 = \frac{\partial f}{\partial s^A} d_\mu s^A$$

4 By an abuse of notation we sometimes identify points in $\mathbf{P}$, $\mathbf{F}$, etc., with their coordinate representation

which can be thought of as exterior differentiation in the $s^{\mathbf{A}}$ variables with $q^{\mathbf{A}}$, r as parameters. Similarly, relative 2-forms may be thought of as 2-forms with respect to $s^{\mathbf{A}}$ depending parametrically on $q^{\mathbf{A}}, r$. We may perform a similar analysis on $\mathbb{F}_1^I$ using coordinates

$$Q^{\mathbf{A}} := (iz^{\mathbf{A}\dot{\mathbf{A}}}\pi_{\dot{\mathbf{A}}})/\pi_{1'}, \quad R := \pi_{0'}/\pi_{1'}$$
$$S^{\mathbf{A}} := -z_{1'}^{\mathbf{A}} = -z^{\mathbf{A}0'}.$$

Thus, on $\mathbb{F}_1^I$, d_μ is given by

$$d_\mu f = \frac{\partial f}{\partial S^{\mathbf{A}}} d_\mu S^{\mathbf{A}}.$$

On $\mathbb{F}_0^I \cap \mathbb{F}_1^I$ we have

$$S^{\mathbf{A}} = -z_{1'}^{\mathbf{A}} = -z^{\mathbf{A}0'} = -(q^{\mathbf{A}}/i) + (z^{\mathbf{A}1'}\pi_{1'}/\pi_{0'}) = -q^{\mathbf{A}}/i + rs^{\mathbf{A}}$$

and thus our frames $d_\mu s^{\mathbf{A}}$ and $d_\mu S^{\mathbf{A}}$ for Ω_μ^1 over $\mathbb{F}_0^I \cap \mathbb{F}_1^I$ are related by

$$d_\mu S^{\mathbf{A}} = r d_\mu s^{\mathbf{A}}. \tag{1.8}$$

Now recall that $r = \pi_{1'}/\pi_{0'}$ is the transition function for $H \to \mathbb{P}_1$, the hyperplane section bundle, so (1.8) implies that

$$\Omega_\mu^1|_{\mathbb{F}^I} \simeq \mathcal{O}_{\mathbb{F}}(1) \oplus \mathcal{O}_{\mathbb{F}}(1)|_{\mathbb{F}^I}. \tag{1.9}$$

Similarly one obtains

$$d_\mu S \wedge d_\mu T = r^2 d_\mu s \wedge d_\mu t$$

from which it follows that

$$\Omega_\mu^2|_{\mathbb{F}^I} \simeq \mathcal{O}_{\mathbb{F}}(2)|_{\mathbb{F}^I}. \tag{1.10}$$

Remark. The isomorphisms (1.9) and (1.10) are not valid on all of $\mathbb{F}$. In the case of (1.9) note that a fibre Y of the fibration $\mathbb{F} \to \mathbb{P}$ is isomorphic to $\mathbb{P}_2(\mathbb{C})$ so that Y has non-trivial holomorphic cotangent bundle. But this is exactly what we obtain by restricting Ω_μ^1 as a bundle to Y whereas $\mathcal{O}_{\mathbb{F}}(1) \oplus \mathcal{O}_{\mathbb{F}}(1)$ is trivial over Y. A similar argument holds for (1.10). The reason why these isomorphisms do not extend is really that relation (1.8) involves a spinor index $\mathbf{A}$ which has been ignored in (1.9) and (1.10). This problem will be rectified after first completing the proof of Lemma 1.1.

Proof of Lemma 1.1. In the coordinates introduced above we see that the exactness of (1.3) follows immediately from the exact (deRham) sequence

$$0 \longrightarrow \mathbb{C} \longrightarrow \mathcal{O}_{\mathbb{C}^2} \xrightarrow{d} \Omega^1_{\mathbb{C}^2} \xrightarrow{d} \Omega^2_{\mathbb{C}^2} \longrightarrow 0$$

by considering the relative forms in $\mathbb{C}^5$ as forms in $\mathbb{C}^2$ depending on three additional parameters. □

From (1.8) it follows immediately that

$$\left(\frac{\partial f}{\partial S^{\mathbf{A}}}\right)\pi_{0'} = \left(\frac{\partial f}{\partial S^{\mathbf{A}}}\right)\pi_{1'} \text{ on } \mathbb{F}_0^I \cap \mathbb{F}_1^I. \tag{1.11}$$

Using the chain rule on the formulae for $q^{\mathbf{A}}$, r, and $s^{\mathbf{A}}$ we find

$$\nabla_{\mathbf{A}}^{0'} := -\frac{\partial}{\partial z_{0'}^{\mathbf{A}}} = \frac{\partial}{\partial s^{\mathbf{A}}} - ir\frac{\partial}{\partial q^{\mathbf{A}}}$$

$$\nabla_{\mathbf{A}}^{1'} := -\frac{\partial}{\partial z_{1'}^{\mathbf{A}}} = i\frac{\partial}{\partial q^{\mathbf{A}}}$$

Thus, we may rewrite the left-hand side of (1.11):

$$\left(\frac{\partial}{\partial s^{\mathbf{A}}}\right)\pi_{0'} = (\nabla_{\mathbf{A}}^{0'} + r\nabla_{\mathbf{A}}^{1'})\pi_{0'} = \pi_{0'}\nabla_{\mathbf{A}}^{0'} + \pi_{1'}\nabla_{\mathbf{A}}^{1'} = \pi_{A'}\nabla_{\mathbf{A}}^{A'}.$$

Rewriting the right-hand side of (1.11), of course, produces the same answer. This representation shows directly how the twist appears in (1.9). It comes from the $\pi_{\mathbf{A}}$ which may be regarded as sections of $\mathcal{O}_{\mathbb{P}_1}(1)$. It also shows how the spinor index $\mathbf{A}$ appears and, indeed, globally Ω_μ^1 may be canonically identified with the *spinor sheaf* $\mathcal{O}_{\mathbf{A}}(1)[-1]'$. This will be fully explained in the next section. Over $\mathbb{F}^I$, however, this spinor sheaf can be trivialized appropriately and this identification reduces to (1.9). From the above discussion the following lemma is immediate over $\mathbb{F}^I$ and the global statement will follow from the arguments of Sect. 2 ($[AB]$ means, as usual, skew-symmetrization of the indices).

Lemma 1.2. *The relative deRham sequence* (1.3) *on* $\mathbb{F}$ *may be canonically identified with the sequence of spinor sheaves*

$$\begin{array}{ccccc}
\Omega_\mu^0 & \longrightarrow & \Omega_\mu^1 & \longrightarrow & \Omega_\mu^2 \\
\| & & \| & & \| \\
\mathcal{O} & \longrightarrow & \mathcal{O}_A(1)[-1]' & \longrightarrow & \mathcal{O}_{[AB]}(2)[-2]' \\
\cup & & \cup & & \cup \\
f & \longmapsto & \pi_{A'}\nabla_A^{A'}f, & g_A & \longmapsto \pi_{A'}\nabla_{[A}^{A'}g_{B]}.
\end{array}$$

2. Spinor Sheaves and Direct Image Sheaves

In order to define correctly the massless fields which are the subject of this paper it is necessary to introduce certain natural holomorphic vector-bundles on $\mathbb{M}$. These bundles are the various *conformally weighted spin-bundles* and the fields of interest on $\mathbb{M}$ are best described as sections thereof satisfying differential equations described in terms of the natural differential operators between these spin-bundles. We shall make use of the identification of a holomorphic vector-bundle with the sheaf of its holomorphic sections, often without explicitly saying so. Thus, our first task is to describe the spinor sheaves on $\mathbb{M}$.

In the previous section we referred to $\mathbb{F}$ as the projective primed spin-bundle. The fibre of $\mathbb{F}$ over a point z in $\mathbb{M}$ is the space of rays in z (regarded as a complex 2-plane in $\mathbb{T}$). Hence, $\mathbb{F}$ is the projective version of the vector bundle over $\mathbb{M}$ which associates with each point z in $\mathbb{M}$ the 2-dimensional vector-subspace of $\mathbb{T}$ which z represents. In topology this bundle is called the *universal bundle* on $\mathbb{M}$ (cf. [26], [51]). As warned above we make no distinction between vector-bundles and locally free sheaves and introduce the notation

$$\mathcal{O}_{A'} := \text{the universal bundle on } \mathbb{M},$$

calling this the *dual primed spin-bundle* so that $\mathbb{P}$ is the projective version of $\mathcal{O}_{A'}$. Its dual, the *primed spin-bundle*, will be denoted $\mathcal{O}^{A'}$. It should be emphasized that $\mathcal{O}_{A'}$ is not a direct sum of two line bundles and that A' is thus not a numerical spinor index but an *abstract index* (cf. [31], [40])[5]. This index is meant to serve as a marker which indicates what sections of this holomorphic vector bundle look like in coordinate patches. For example, in terms of such coordinates, $f^{A'} \in \mathcal{O}^{A'}$ is represented by $f^{A'} = (f^{0'}, f^{1'}) \in \mathcal{O} \oplus \mathcal{O}$, the map from $f^{A'}$ to $f^{\mathbf{A}'}$ being achieved by use of the basis elements (local frame) $\varepsilon_{A'}^{\mathbf{B}'} = (\varepsilon_{A'}^{0'}, \varepsilon_{A'}^{1'}) \in \mathcal{O}_{A'} \oplus \mathcal{O}_{A'}$, with $f^{\mathbf{B}'} = f^{A'} \varepsilon_{A'}^{\mathbf{B}'}$, the *repeated index* A' denoting the abstract scalar product between a vector and an element of its dual space. The map back from $f^{\mathbf{A}'}$ to $f^{A'}$ is achieved by the dual basis $\varepsilon_{\mathbf{B}'}^{A'} = (\varepsilon_{0'}^{A'}, \varepsilon_{1'}^{A'}) \in \mathcal{O}^{A'} \oplus \mathcal{O}^{A'}$ according to $f^{A'} = f^{\mathbf{B}'} \varepsilon_{\mathbf{B}'}^{A'}$, the repeated index $\mathbf{B}'$ now indicating numerical summation convention. Correspondingly, the relation between $g_{A'} \in \mathcal{O}_{A'}$ and $g_{\mathbf{A}} \in \mathcal{O} \oplus \mathcal{O}$ is given by $g_{\mathbf{B}'} = g_{A'} \varepsilon_{\mathbf{B}'}^{A'}$ and $g_{A'} = g_{\mathbf{B}'} \varepsilon_{A'}^{\mathbf{B}'}$ and the scalar product between $g_{A'}$ and $f^{A'}$ is $g_{A'} f^{A'} = g_{\mathbf{A}} f^{\mathbf{A}'} = g_{0'} f^{0'} + g_{1'} f^{1'}$. One can choose bases $\varepsilon_{A'}^{\mathbf{B}'}$ and $\varepsilon_{B'}^{\mathbf{A}'}$ which effect a trivialization of $\mathcal{O}_{A'}$ and $\mathcal{O}^{A'}$ over the subspace $\mathbb{M}^I$ of $\mathbb{M}$ (by using the coordinates given in (1.7), for instance).

To define the unprimed spin-bundle over $\mathbb{M}$ we observe that $\mathcal{O}_{A'}$ is, by definition, canonically embedded in $\mathcal{O}^{\alpha}$, the trivial bundle over $\mathbb{M}$ whose fibre at each point is twistor-space $\mathbb{T}^{\alpha}$. The *unprimed spin-bundle* $\mathcal{O}^{A}$ is defined to be the quotient (or *complement*). In other words, we have the short exact sequence[6]

$$\mathbb{M};\ 0 \longrightarrow \mathcal{O}_{A'} \longrightarrow \mathcal{O}^{\alpha} \longrightarrow \mathcal{O}^{A} \longrightarrow 0 \tag{2.1}$$

(cf. the definition of "local twistors" e.g. Penrose and MacCallum [32]). The dual unprimed spin-bundle is denoted $\mathcal{O}_{A}$, and each of $\mathcal{O}^{A}$ and $\mathcal{O}_{A}$ may be trivialized over $\mathbb{M}^I$ by means of bases, e.g.

$$\mathbb{M}^I;\ \mathcal{O} \oplus \mathcal{O} \ni (f_0, f_1) = f_{\mathbf{A}} \mapsto f_A = f_{\mathbf{B}} \varepsilon_A^{\mathbf{B}} \in \mathcal{O}_{\mathbf{A}}.$$

Taking the dual of (2.1) gives the exact sequence

$$\mathbb{M};\ 0 \longrightarrow \mathcal{O}_{A} \longrightarrow \mathcal{O}_{\alpha} \longrightarrow \mathcal{O}^{A'} \longrightarrow 0 \tag{2.2}$$

The similarity between (2.1) and (2.2) bears further investigation. Let $\mathbb{M}^*$ denote the Grassmannian of 2-planes in $\mathbb{T}_{\alpha}$, the dual twistor space. Then there is a canonical isomorphism

$$\begin{array}{ccc} i: \mathbb{M} & \xrightarrow{\cong} & \mathbb{M}^* \\ \cup & & \cup \\ z & \longmapsto & z^{\perp} := \{W_{\alpha} \in \mathbb{T}_{\alpha} : Z^{\alpha} W_{\alpha} = 0, \forall Z^{\alpha} \in z\}. \end{array}$$

5 We adopt the "abstract index" notation (Penrose [31]). Abstract indices are non-numerical and serve only as organizational markers enabling the basically coordinate-free operations of contraction, symmetrization, index permutation, etc. to be expressed in a transparent yet frame-independent way. The abstract indices are simply labels for bundles or sections of a given bundle. Italic and Greek letters will be used in this abstract way while Sans serif letters (**A**, **B**, ...) will be used for the corresponding "normal" indices (i.e. to represent the components in some particular frame). Italic lower case indices will be used for tensor indices, italic upper case for spinor indices, and Greek will be used for abstract twistor indices. Sans serif lower case indices **a**, **b**, **c**, ... will take values in the range {0, 1, 2, 3} whilst Sans serif upper case indices **A**, **B**, **C**, ... will take values in the range {0, 1}.

6 From now on, a space X at the left of a diagram (such as $\mathbb{M}$ in equation (2.1)) indicates that X is the base-space of the various sheaves involved

We can now go through exactly the same process to describe various bundles on $\mathbb{M}^*$. Since $\mathbb{T}_\alpha$ has coordinates $(\eta_A, \xi^{A'})$, it is natural to denote the universal bundle on $\mathbb{M}^*$ by $\mathcal{O}_A$ and define $\mathcal{O}^{A'}$ by means of a sequence which formally looks just the same as (2.2). This abuse of notation is justified by the observation that there are canonical isomorphisms

$$z \cong (\mathbb{T}_{\alpha'}/z^\perp)^*, \quad \mathbb{T}^\alpha/z \cong (z^\perp)^*, \text{ etc.},$$

giving rise to canonical isomorphisms

$$(\mathcal{O}_{A'} \text{ on } \mathbb{M}) \cong i^*(\mathcal{O}_{A'} \text{ on } \mathbb{M}^*), (\mathcal{O}^A \text{ on } \mathbb{M}) \cong i^*(\mathcal{O}^A \text{ on } \mathbb{M}^*), \text{ etc.}$$

But now observe that any isomorphism $p : \mathbb{T}^\alpha \xrightarrow{\simeq} \mathbb{T}_\alpha$ gives rise to an isomorphism

$$p : \mathbb{M} \xrightarrow{\simeq} \mathbb{M}^*$$

under which universal bundles naturally correspond i.e. $\mathcal{O}_{A'} = p^*\mathcal{O}_A$. Hence we conclude that any isomorphism of $\mathbb{T}^\alpha$ with $\mathbb{T}_\alpha$ naturally extends to an isomorphism of (2.1) with (2.2). Thus $\mathcal{O}_{A'}$ and $\mathcal{O}_A$ are on an equal footing when $\mathbb{M}$ is regarded as an abstract complex manifold rather than a specific Grassmannian. Choosing one of these bundles to be called the primed spin-bundle is, in this sense, arbitrary but it corresponds to making a choice of *complex orientation* for $\mathbb{M}$.

Remark. We want to distinguish in this paper between "isomorphism" and "canonical isomorphism," both used in the above discussion. An intuitive notion of what is canonical ("God-given"), and what is not is common among mathematicians, but it is worthwhile to clarify the point. In general, the meaning of the word depends on the context. The idea is that something is canonical if it transforms correctly when the whole system is subjected to a morphism of the category in which one is working (often not explicitly mentioned). In this paper, our basic given data is the 4-dimensional vector space $\mathbb{T}$, and a canonical isomorphism for us will mean an isomorphism of two quantities dependent upon $\mathbb{T}$ but which is preserved under the action of $GL(\mathbb{T})$ on $\mathbb{T}$. In the discussion above, we see that $i : \mathbb{M} \xrightarrow{\cong} \mathbb{M}^*$ is a canonical isomorphism whereas $p : \mathbb{M} \xrightarrow{\simeq} \mathbb{M}^*$ is not. We use "$\cong$" to denote canonical isomorphism, and "$\simeq$" to refer to an isomorphism, not necessarily canonical (usually where some choice has been made, e.g. $\mathbb{M} \simeq G_{2,4}(\mathbb{C})$ corresponds to a choice of basis in $\mathbb{T}$).

From these two basic bundles we can generate lots more by applying standard tensor operations. The following notation is natural and essentially self-explanatory[7]:

7 Here we mean tensor products as $\mathbb{C}$-vector bundles, i.e. the fibre-wise tensor product of the vector bundle fibres as complex vector spaces. It may be remarked, also, that the abstract index notation entails that $\mathcal{O}_{A'}$ and $\mathcal{O}_{B'}$ are canonically isomorphic copies of one another, but not, strictly speaking, the same space. The same applies to $\mathcal{O}_{C'}$, etc. The identification of $\mathcal{O}_{A'B'}$ with $\mathcal{O}_{A'} \otimes \mathcal{O}_{B'}$ rather than with $\mathcal{O}_{B'} \otimes \mathcal{O}_{A'}$ involves the notion of an alphabetical ordering for the abstract indices. In fact, $\mathcal{O}_{B'A'} \cong \mathcal{O}_{A'B'} = \mathcal{O}_{A'} \otimes \mathcal{O}_{B'} \cong \mathcal{O}_{B'} \otimes \mathcal{O}_{A'} \cong \mathcal{O}_{A'C'} = \mathcal{O}_{A'} \otimes \mathcal{O}_{C'}$, etc. (cf. [31], [40])

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

$$\mathcal{O}_{A'B'} := \mathcal{O}_{A'} \otimes \mathcal{O}_{B'}, \quad \mathcal{O}_{AA'} := \mathcal{O}_A \otimes \mathcal{O}_{A'}$$
$$\mathcal{O}_{(A'B')} := \mathcal{O}_{A'} \odot \mathcal{O}_{B'} \text{ where } \odot \text{ means symmetric product}$$
$$\mathcal{O}_{[A'B']} := \mathcal{O}_{A'} \wedge \mathcal{O}_{B'} = \det(\mathcal{O}_{A'}), \quad \mathcal{O}_A^{A'} := \mathcal{O}_A \otimes \mathcal{O}^{A'}, \text{ etc., etc.}$$

Of particular interest are the two natural line bundles $\mathcal{O}_{[AB]}$ and $\mathcal{O}_{[A'B']}$. We introduce an alternative notation:

$$\mathcal{O}_{[AB]} := \mathcal{O}[-1] \quad \text{and} \quad \mathcal{O}_{[A'B']} := \mathcal{O}[-1]'.$$

This is in line with the notation of $\mathcal{O}(-1)$ for the universal bundle on a projective space. The number -1 in the square brackets may be regarded, with suitable normalization, as the first Chern class of these bundles. Just as for projective space, $\mathcal{O}[1]$ (resp. $\mathcal{O}[1]'$) will denote the dual of $\mathcal{O}[-1]$ (resp. $\mathcal{O}[-1]'$) and, more generally,

$$\mathcal{O}[k] := \{\mathcal{O}[1]\}^k \quad \text{and} \quad \mathcal{O}[k]' := \{\mathcal{O}[1]'\}^k.$$

We can now rewrite, for example,

$$\mathbb{M};\ \mathcal{O}_{A[A'B']} = \mathcal{O}_A \otimes \mathcal{O}_{[A'B']} = \mathcal{O}_A \otimes \mathcal{O}[-1]' := \mathcal{O}_A[-1]',$$

and write $\mathcal{O}_{(AB)}[-1][2]'$ to mean $\mathcal{O}_{(AB)} \otimes \mathcal{O}[-1] \otimes \mathcal{O}[2]'$, etc. Our notation has some redundancies. For example, there is the canonical isomorphism

$$\mathbb{M};\ \mathcal{O}_{[C'D']}[1]' := \mathcal{O}_{[C'D']} \otimes \mathcal{O}^{[A'B']} \cong \mathcal{O}.$$

Under this isomorphism we shall denote by $\varepsilon_{A'B'}$ the section of $\mathcal{O}_{[A'B']}[1]'$ corresponding to $1 \in \mathcal{O}$. Similarly, we have

$$\varepsilon_{AB} \in \mathcal{O}_{[AB]}[1], \varepsilon^{A'B'} \in \mathcal{O}^{[A'B']}[-1]', \varepsilon^{AB} \in \mathcal{O}^{[AB]}[-1].$$

These are to be regarded as the usual conformally weighted ε's used for raising and lowering spinor indices. Thus, we see that the numbers in the square brackets may be regarded as *conformal weights*, with $\varepsilon_{A'B'}$, for example, having conformal weight 1. This entails, roughly speaking, that a metric g_{ab} is replaced by $\lambda\lambda' g_{ab}$ when $\varepsilon_{A'B'}$ is replaced by $\lambda'\varepsilon_{A'B'}$ and ε_{AB} by $\lambda\varepsilon_{AB}$ (whereas $\varepsilon^{A'B'}$ must be replaced by $(\lambda')^{-1}\varepsilon^{A'B'}$ and ε^{AB} by $\lambda^{-1}\varepsilon^{AB}$). Moreover, we observe that these ε's really can be used for raising and lowering spinor indices on sheaves. For example:

$$\begin{array}{lccccc} & \mathcal{O}^{A'} & \cong & \mathcal{O}^{A'} \otimes \mathcal{O} & \cong & \mathcal{O}^{A'} \otimes \mathcal{O}_{[A'B']}[1]' \cong \mathcal{O}_{B'}[1]' \\ \mathbb{M}; & \cup & & \cup & & \cup \qquad\qquad\qquad\qquad \cup \\ & f^{A'} & \longleftrightarrow & f^{A'} \otimes 1 & \longleftrightarrow & f^{A'} \otimes \varepsilon_{A'B'} \longleftrightarrow f^{A'}\varepsilon_{A'B'} \end{array}$$

All these bundles may be pulled back to $\mathbb{F}$ and we will often use the same notation to denote the pullbacks. Thus, we write $\mathcal{O}_A$ instead of $\nu^*\mathcal{O}_A$ if it is clear that we are discussing bundles on $\mathbb{F}$ rather than $\mathbb{M}$. Also, we can pull the bundles $\mathcal{O}(k)$ on $\mathbb{P}$ back to $\mathbb{F}$. Again we make no notational distinction and naturally write

$$\mathcal{O}_A(1)[-1]' \text{ instead of } \nu^*\mathcal{O}_A[-1]' \otimes \mu^*\mathcal{O}(1) \text{ etc.}$$

This makes sense of all the sheaves in Lemma 1.2 and our next job is to make

sense of the differential operators. We claim that, on $\mathbb{M}$,

$$\nabla_{AA'} : \mathcal{O} \longrightarrow \mathcal{O}_{AA'}$$

is a well-defined intrinsic differential operator. Indeed, we claim that it may be canonically identified with the exterior derivative

$$d : \mathcal{O} \longrightarrow \Omega^1.$$

To make full use of the abstract index notation [31], we should also write this as

$$\nabla_a : \mathcal{O} \longrightarrow \mathcal{O}_a.$$

In local coordinates d and $\nabla_{AA'}$ certainly look identical. The problem is to see that Ω^1 is canonically isomorphic to $\mathcal{O}_{AA'}$. Equivalently, the problem is to show that the holomorphic tangent bundle can be identified naturally with $\mathcal{O}^{AA'}$. By definition,

$$\mathbb{M};\ \mathcal{O}^{AA'} = \mathcal{O}^A \otimes \mathcal{O}^{A'} = \mathcal{O}^A \otimes (\mathcal{O}_{A'})^* \cong \mathrm{Hom}(\mathcal{O}_{A'}, \mathcal{O}^A).$$

It is well known, however, that the tangent bundle of a Grassmannian manifold is isomorphic to the bundle of homomorphisms from the universal bundle to its complement (cf. [26]). Other differential operators are defined by using ε's to raise indices. For example, $\nabla_A^{A'}$ is defined as the composition

$$\mathbb{M};\ \mathcal{O} \xrightarrow{\nabla_{AB'}} \mathcal{O}_{AB'} \xrightarrow{\varepsilon^{A'B'}} \mathcal{O}_A^{A'}[-1]'.$$

The exterior derivative $d : \Omega^1 \to \Omega^2$ may also be written out in terms of spinor sheaves. Ω^2 splits into its self-dual and anti-self-dual parts: $\Omega^2 = \Omega^2_+ \oplus \Omega^2_-$. Decomposing a 2-form F_{ab} as $F^+_{ab} + F^-_{ab}$ is easily written in spinor notation (cf. [52], [40]):

$$F_{ab} = F_{AA'BB'} = \phi_{A'B'}\varepsilon_{AB} + \psi_{AB}\varepsilon_{A'B'},$$

where $\phi_{A'B'}$ and ψ_{AB} are symmetric. Alternatively, in terms of bundles,

$$\begin{aligned}\mathbb{M};\ \Omega^2 \cong \mathcal{C}_{[ab]} \cong \Omega^2_+ \oplus \Omega^2_- &\cong \mathcal{O}_{(A'B')[AB]} \oplus \mathcal{O}_{(AB)[A'B']} \\ &= \mathcal{O}_{(A'B')}[-1] \oplus \mathcal{O}_{(AB)}[-1]'\end{aligned}$$

and the exterior derivative may be written

$$\begin{array}{rccc} \mathbb{M}; & \mathcal{O}_{AA'} & \longrightarrow & \mathcal{O}_{(A'B')}[-1] \oplus \mathcal{O}_{(AB)}[-1]' \\ & \cup & & \cup \\ & \phi_{AA'} & \longmapsto & \nabla^A_{(B'}\phi_{A')A} \quad + \nabla^{A'}_{(B}\phi_{A)A'} \end{array}$$

More generally, the whole deRham sequence may be written:

$$\begin{array}{cccccccccc} \Omega^0 & \longrightarrow & \Omega^1 & \longrightarrow & \Omega^2 & \longrightarrow & \Omega^3 & \longrightarrow & \Omega^4 \\ \wr\| & & \wr\| & & \wr\| & & \wr\| & & \wr\| \\ \mathcal{O} & \xrightarrow{\nabla_{AA'}} & \mathcal{O}_{AA'} & \xrightarrow{\begin{matrix}\nabla^{A'}_B \\ \nabla^{A'}_B\end{matrix}} & \begin{matrix}\mathcal{O}_{(A'B')}[-1] \\ \oplus \\ \mathcal{O}_{(AB)}[-1]'\end{matrix} & \xrightarrow{\begin{matrix}\nabla^{A'}_B \\ i\nabla^A_B\end{matrix}} & \mathcal{O}_{BB'}[-1][-1]' & \xrightarrow{\nabla_{BB'}} & \mathcal{O}[-2][-2]' \end{array} \qquad (2.3)$$

These operators also act between other pairs of spinor sheaves as we will see when we discuss direct images.

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

To make sense of the differential operator

$$\mathbb{F};\ \pi_{A'}\nabla_A^{A'}:\mathcal{O}\longrightarrow\mathcal{O}_A(1)[-1]'$$

which appears in Lemma 1.2 we first observe that we can invariantly define a bundle homomorphism

$$\mathbb{F};\ \pi_{A'}:\mathcal{O}^{A'}\longrightarrow\mathcal{O}(1).$$

Indeed, it is equivalent to define the dual

$$\mathbb{F};\ {}^t\pi_{A'}:\mathcal{O}(-1)\longrightarrow\mathcal{O}_{A'}$$

and geometrically this can be seen as follows. By definition, the fibres of $\mathcal{O}(-1)$ and $\mathcal{O}_{A'}$ over $(L_1, L_2)\in\mathbb{F}$ are precisely L_1 and L_2 respectively so that ${}^t\pi_{A'}$ may be defined over (L_1, L_2) as the inclusion $L_1\hookrightarrow L_2$. One could hope, perhaps, that $\pi_{A'}\nabla_A^{A'}$ might be defined as the composition

$$\mathbb{F};\ \mathcal{O}\xrightarrow{\nabla_A^{A'}}\mathcal{O}_A^{A'}[-1]'\xrightarrow{\pi_{A'}}\mathcal{O}_A(1)[-1]'$$

but this is not the case as $\nabla_{AA'}$ does not make sense on $\mathbb{F}$. If, however, we choose a specific trivialization of $\mathbb{F}$ over a neighbourhood of a point in $\mathbb{M}$, then, of course, $\nabla_{AA'}$ does make sense as differentiation in the base direction. We claim that then $\pi_{A'}\nabla_A^{A'}$ is independent of this choice of trivialization. Certainly this should be the case since $\pi_{A'}\nabla_A^{A'}$ is supposed to be d_μ which is, by definition, differentiation along the fibres of μ and these fibres are mapped isomorphically by ν to the α-planes in $\mathbb{M}$ [52]. Thus we can think of d_μ as first restricting to a fibre of μ, then identifying this fibre with an α-plane in $\mathbb{M}$, and finally effecting the exterior derivative on this α-plane. This last step can certainly be expressed using just $\nabla_{AA'}$, the exterior derivative on $\mathbb{M}$. We can therefore justify our claim by showing that the map from the cotangent bundle of $\mathbb{M}$ restricted to some α-plane to the cotangent bundle of the α-plane itself is given by $\pi^{A'}$. This is easy to see geometrically from the definitions of the spaces involved.

It is, perhaps, worth while to give an alternative way of describing d_μ as $\pi_{A'}\nabla_A^{A'}$ more directly. If we choose local coordinates $(x^{AA'}, \pi_{A'})$ on $\mathbb{F}$ then $\pi^{A'}\nabla_{AA'}$ makes perfectly good sense and, by our calculations of Sect. 1, does indeed represent d_μ. Actually, it is easy to see that if f is a function which is constant on the fibres of μ i.e. f depends on $\omega^A(:=ix^{AA'}\pi_{A'})$ and $\pi_{A'}$ alone, then f is annihilated by $\pi^{A'}\nabla_{AA'}$:

$$\pi^{A'}\nabla_{AA'}f(\omega^B,\pi_{B'})=\pi^{A'}\left(i\pi_{A'}\frac{\partial f}{\partial\omega_A}(\omega^B,\pi_{B'})\right)=0.$$

The invariance of $\pi^{A'}\nabla_{AA'}$ may then be checked directly, by changing from one local coordinate system to another.

It may be helpful, at this stage, to point out precisely the lack of invariance that can arise in expressions involving the operator $\nabla_{AA'}$. Each of the trivializations of $\mathbb{F}$ that we have been considering is associated with a particular flat connection $\nabla_{AA'}$ on a portion of $\mathbb{M}$ (e.g. on the $\mathbb{M}^I$ of Sect. 1). The non-invariance that may arise in the use of $\nabla_{AA'}$ can be seen in explicit transformation formulae relating one such connection to another. It is helpful, moreover, to examine such trans-

formation formulae in the broader context of general torsion-free connections on $\mathbb{M}$ that are compatible with the (complex) conformal structure of $\mathbb{M}$. We require $\nabla_{AA'}$ to act on spinor fields on some open $W \subset \mathbb{M}$ and, correspondingly, to provide a bundle connection on $W' \subset \mathbb{F}$. We suppose g_{ab} and $\hat{g}_{ab}$ are two (not necessarily flat) complex metrics on W compatible with the given conformal structure, with spinor representations given as

$$g_{AA'BB'} = \varepsilon_{AB}\varepsilon_{A'B'}, \hat{g}_{AA'BB'} = \hat{\varepsilon}_{AB}\hat{\varepsilon}_{A'B'}$$

where

$$\hat{g}_{ab} = \Omega\tilde{\Omega}g_{ab}, \hat{\varepsilon}_{AB} = \Omega\varepsilon_{AB}, \hat{\varepsilon}_{A'B'} = \tilde{\Omega}\varepsilon_{A'B'},$$

the Ω and $\tilde{\Omega}$ being arbitrary non-vanishing holomorphic scalar functions on W. For spinor fields on $\mathbb{M}$, the relevant formulae are essentially given in [31], [32], [40]:

$$(\hat{\nabla}_{AA'} - \nabla_{AA'})\chi_{B\cdots F'\cdots}^{P\cdots S'\cdots} = \varepsilon_A{}^P\Upsilon_{XA'}\chi_{B\cdots F'\cdots}^{X\cdots S'\cdots} + \ldots + \varepsilon_{A'}{}^{S'}\Upsilon_{AX'}\chi_{B\cdots F'\cdots}^{P\cdots X'\cdots} + \ldots$$
$$- \Upsilon_{BA'}\chi_{A\cdots F'\cdots}^{P\cdots S'\cdots} - \ldots - \Upsilon_{AF'}\chi_{B\cdots A'\cdots}^{P\cdots S'\cdots} - \ldots + k\Pi_{AA'}\chi_{B\cdots F'\cdots}^{P\cdots S'\cdots} \quad (2.4)$$

where

k = (no. of upper unprimed and lower primed indices)
$-$ (no. of upper primed and lower unprimed indices)

and where

$$2\Upsilon_{AA'} = \Omega^{-1}\nabla_{AA'}\Omega + \tilde{\Omega}^{-1}\nabla_{AA'}\tilde{\Omega}, \quad 4\Pi_{AA'} = \Omega^{-1}\nabla_{AA'}\Omega - \tilde{\Omega}^{-1}\nabla_{AA'}\tilde{\Omega},$$

these formulae being uniquely characterized by the additivity, Leibniz and torsion-free properties, and the conditions

$$\nabla_{AA'}\varepsilon_{BC} = 0 = \hat{\nabla}_{AA'}\hat{\varepsilon}_{BC}$$
$$\nabla_{AA'}\varepsilon_{B'C'} = 0 = \hat{\nabla}_{AA'}\hat{\varepsilon}_{B'C'}.$$

The formula (2.4) is useful for checking the invariance of many of the operations of this paper, for example, the massless field equations (of appropriate weights) considered in Sect. 5–Sect. 9, where a quantity $\chi_{\cdots}^{\cdots}$ of weights $[p][q]'$, when acted upon by $\hat{\nabla}_{AA'}$, is replaced by

$$\hat{\chi}_{\cdots}^{\cdots} = \Omega^p\tilde{\Omega}^q\chi_{\cdots}^{\cdots}.$$

In fact, the difference between the primed and unprimed weights does not seem to play much of a role here, as it turns out that a $GL(\mathbb{T})$ transformation always produces an Ω and $\tilde{\Omega}$ which are constant multiples of one another, so here $\Pi_{AA'} = 0$.

For a scalar field f on $\mathbb{F}$ we find, with $\Pi_{AA'} = 0$,

$$(\hat{\nabla}_{AA'} - \nabla_{AA'})f = \pi_{A'}\Upsilon_{AC'}\frac{\partial f}{\partial \pi_{C'}}, \quad (2.5)$$

where the fibre-wise derivative $\partial f/\partial\pi_{C'}$ has an obvious invariant meaning. The general form of the right-hand side of (2.5), i.e. with $\pi_{A'}\Upsilon_{AC'}$ replaced by some $Q_{AA'C'}$ independent of f, follows from the fact that each of $\hat{\nabla}_{AA'}$ and $\nabla_{AA'}$ gives simply d on $\mathbb{M}$. The particular expression given in (2.5) follows from (2.4), for

example, by taking

$$f = f^{P'\cdots R'}\pi_{P'}\cdots\pi_{R'},$$

where $f^{P'\cdots R'}$ is a spinor field on $\mathbb{M}$. We can treat expressions with both spinor indices and π-dependence by adding the appropriate term (2.5) to the expression (2.4).

The invariance of $\pi^{A'}\nabla_{AA'}$ when acting on unweighted scalar quantities on $\mathbb{F}$ is now immediately obvious.

Note that we may use ε^{AB} to rewrite the relative deRham sequence in Lemma 1.2 as

$$\begin{array}{lccccc} & \Omega^0_\mu & \xrightarrow{d_*} & \Omega^1_\mu & \xrightarrow{d_*} & \Omega^2_\mu \\ \mathbb{F}: & \| & & \| & & \| \\ & \mathcal{O} & \xrightarrow{\pi_A\nabla^A_{A'}} & \mathcal{O}_A(1)[-1]' & \xrightarrow{\pi_{A'}\nabla^{AA'}} & \mathcal{O}(2)[-1][-2]' \end{array}$$

or alternatively (with a change of sign which we suppress)

$$\mathbb{F};\ \mathcal{O} \xrightarrow{\pi_{A'}\nabla^{AA'}} \mathcal{O}^A(1)[-1][-1]' \xrightarrow{\pi_{A'}\nabla^{A'}_A} \mathcal{O}(2)[-1][-2]'.$$

We now want to consider sheaves on $\mathbb{M}$ derived from these sheaves Ω^p_μ and more generally from $\Omega^p_\mu(V)$ for a holomorphic vector bundle V over a portion of $\mathbb{P}$. We shall use the notion of *direct image sheaves* (cf. Bredon [5], Godement [14]) which are defined as follows. If $f: X \to Y$ is a continuous mapping of topological spaces, and $\mathscr{F}$ is a sheaf of Abelian groups on X then we define the direct images $f^q_*\mathscr{F}$ on Y by means of the presheaf

$$U \longmapsto H^q(f^{-1}(U), \mathscr{F}), \quad \text{for } U^{\text{open}} \subseteq Y$$

with the obvious restriction mapping. Thus, we see that the stalk of $f^q_*\mathscr{F}$ at $y \in Y$ is given by

$$(f^q_*\mathscr{F})_y = \varinjlim_{U \ni y} H^q(f^{-1}(U), \mathscr{F}),$$

i.e. the "cohomology along the fibre" of the sheaf $\mathscr{F}$. In the most natural case when $V = H^n$ the direct images $\nu^q_*(\Omega^p_\mu(n))$ turn out to be various spin-bundles on $\mathbb{M}$ as already defined in this section. This identification will be deduced from Lemma 1.2 and the following (letting $\nu_* := \nu^0_*$).

Proposition 2.1. *The direct images $\nu^q_*\mathcal{O}(k)$ of the sheaves $\mathcal{O}(k)$ on $\mathbb{F}$ may be canonically identified as:*

$\mathbb{M}$;
a) if $n \geqq 0$: $\nu_*\mathcal{O}(n) \cong \mathcal{O}^{(A'B'\cdots D')}$ (n indices)
$\nu^q_*\mathcal{O}(n) = 0 \quad \forall q \geqq 1$
b) $\nu^q_*\mathcal{O}(-1) = 0 \quad \forall q$
c) if $n \geqq 0$: $\nu^1_*\mathcal{O}(-n-2) \cong \mathcal{O}_{(A'B'\cdots D')}[-1]'$ (n indices)
$\nu^q_*\mathcal{O}(-n-2) = 0 \quad \forall q \neq 1.$

Proof. By definition, our task is to compute $H^q(\nu^{-1}(U), \mathcal{O}(k))$ for open sets U in

$\mathbb{M}$ but since there are arbitrarily small Stein open sets[8] it suffices to restrict our attention to the case U Stein and where $\nu^{-1}(U)$ may be trivialized: $U' := \nu^{-1}(U) \simeq U \times \mathbb{P}_1$. Choose coordinates $[\pi_{0'}, \pi_{1'}]$ for $\mathbb{P}_1$. We first look at cohomology on $\mathbb{P}_1$ and then consider the additional parameters in U. If we cover $\mathbb{P}_1$ with

$$V_0 := \{[\pi_{0'}, \pi_{1'}] \in \mathbb{P}_1 : \pi_{0'} \neq 0\}$$
$$V_1 := \{[\pi_{0'}, \pi_{1'}] \in \mathbb{P}_1 : \pi_{1'} \neq 0\},$$

then $V_0 \cong V_1 \cong \mathbb{C}$ which is Stein. We may therefore use the Mayer–Vietoris sequence (cf. Bredon [5])

$$0 \longrightarrow \Gamma(\mathbb{P}_1, \mathcal{O}(k)) \longrightarrow \Gamma(V_0, \mathcal{O}(k)) \oplus \Gamma(V_1, \mathcal{O}(k) \longrightarrow \Gamma(V_0 \cap V_1, \mathcal{O}(k))$$
$$\longrightarrow H^1(\mathbb{P}_1, \mathcal{O}(k)) \longrightarrow H^1(V_0, \mathcal{O}(k)) \oplus H^1(V_1, \mathcal{O}(k)) \longrightarrow \cdots$$
$$\| \qquad\qquad \|$$
$$0 \qquad\qquad 0 : (\text{since } V_0, V_1 \text{ are Stein}),$$

to compute $H^q(\mathbb{P}_1, \mathcal{O}(k))$. The computation consists of expanding elements of $\Gamma(V_0, \mathcal{O}(k)), \Gamma(V_1, \mathcal{O}(k))$, and $\Gamma(V_0 \cap V_1, \mathcal{O}(k))$ in Laurent series and comparing coefficients (cf. Griffiths and Adams [16]). Now since $U' \simeq U \times \mathbb{P}_1$ we may cover U' by two Stein subsets

$$U_1' \cong U \times V_0$$
$$U_1' \cong U \times V_1$$

and proceed to compute $H^q(U', \mathcal{O}(k))$ in terms of this Leray covering by exactly the same technique of Laurent series expansion. The only difference is that the coefficients of the expansions will be holomorphic functions on U. This follows since the coefficients may be defined by means of integral formulae on $\mathbb{P}_1$. Thus we obtain easily

$$H^q(U, \mathcal{O}(k)) = \{\text{holomorphic functions} : U \longrightarrow H^q(\mathbb{P}_1, \mathcal{O}(k))\}. \qquad (2.6)$$

The vanishing statements in the proposition now follow from (2.6) and the well-known vanishing theorem for $\mathbb{P}_1$ (cf. Wells [51])

$$\begin{aligned} H^q(\mathbb{P}_1, \mathcal{O}(n)) &= 0 \quad \text{for } n \geqq 0, \quad q \geqq 1 \\ H^q(\mathbb{P}_1, \mathcal{O}(-1)) &= 0 \quad \text{for all } q \\ H^q(\mathbb{P}_1, \mathcal{O}(-n-2)) &= 0 \quad \text{for } n \geqq 0, \quad q \neq 1. \end{aligned} \qquad (2.7)$$

Alternatively, these statements follow from the Laurent series expansion argument. It is also clear from (2.6) that the direct images $\nu_*^q \mathcal{O}(k)$ in general represent vector bundles whose fibre at $z \in \mathbb{M}$ may be canonically identified with $H^q(\nu^{-1}(z), \mathcal{O}_{\nu^{-1}(z)}(k))$. The problem is to see how these fibres fit together. The simplest case is a), to identify $\nu_* \mathcal{O}(n)$. Clearly $\nu_* \mathcal{O} = \mathcal{O}$ because the fibres of ν are compact so admit only constant holomorphic functions. Thus, let us consider $\nu_* \mathcal{O}(1)$. For any point

8 See Gunning and Rossi [18] and Hörmander [20] for the theory of holomorphic functions on open complex manifolds, in particular Stein manifolds

$z \in \mathbb{M}$ the fibre $\nu^{-1}(z)$ is by definition the projective space $\mathbb{P}(\mathbb{C}_{A'})$, where $\mathbb{C}_{A'}$ is the fibre over z of the primed spin-bundle $\mathcal{O}_{A'}$. Now, $\mathcal{O}_{\mathbb{P}(\mathbb{C}_{A'})}(1)$ may be defined invariantly as the line-bundle which associates to every line in $\mathbb{C}_{A'}$ (i.e. to every point in $\mathbb{P}(\mathbb{C}_{A'})$) its dual (as a complex vector space). Every element of $\mathbb{C}^{A'}$, the dual of $\mathbb{C}_{A'}$, may be restricted to the lines in $\mathbb{C}_{A'}$ thus giving a section of $\mathcal{O}_{\mathbb{P}(\mathbb{C}_{A'})}(1)$. Thus, we have a canonical homomorphism

$$\mathbb{C}^{A'} \longrightarrow \Gamma(\mathbb{P}(\mathbb{C}_{A'}), \mathcal{O}_{\mathbb{P}(\mathbb{C}_{A'})}(1))$$

and the Laurent series argument shows easily that this is an isomorphism. Hence, we can identify $H^0(\nu^{-1}(z), \mathcal{O}_{\nu^{-1}(z)}(1))$ with $\mathbb{C}^{A'}$, the fibre of $\mathcal{O}^{A'}$ over z, in a completely coordinate free manner. It follows that $\nu_* \mathcal{O}(1) \cong \mathcal{O}^{A'}$ as required. The general result that $\nu_* \mathcal{O}(n) \cong \mathcal{O}^{(A'B'\cdots D')}$ follows in the same sort of way. Now we come to case c), to identify $\nu_*^1 \mathcal{O}(-n-2)$. This is not quite so simple as the case $\nu_*^1 \mathcal{O}(-2)$ already shows. The problem is that, although $H^1(\mathbb{P}_1, \mathcal{O}(-2))$ is isomorphic to $\mathbb{C}$, this isomorphism is not canonical but depends on a choice of coordinates for $\mathbb{P}_1$, for example. What is true is that there is a canonical isomorphism $H^1(\mathbb{P}_1, \Omega^1) \cong \mathbb{C}$ (given by integration: Serre duality [43]) and a noncanonical isomorphism: $\mathcal{O}(-2) \simeq \Omega^1$. The isomorphism of Serre duality may be given as follows. Identifying $\mathbb{P}_1$ with $\mathbb{P}(\mathbb{C}_{A'})$ as before we have the exact sequence

$$\mathbb{P}_1 ; \; 0 \longrightarrow \Omega^1 \longrightarrow \mathcal{O}^{A'}(-1) \xrightarrow{\pi_{A'}} \mathcal{O} \longrightarrow 0 \tag{2.8}$$

where $\pi_{A'} : \mathcal{O}^{A'} \to \mathcal{O}(1)$ is defined as the dual of ${}^t\pi_{A'} : \mathcal{O}(-1) \to \mathcal{O}_{A'}$ whose definition is coordinate free (as $\mathcal{O}(-1)$ is the universal bundle which, by definition, is a sub-bundle of $\mathcal{O}_{A'}$, the trivial bundle with fibre $\mathbb{C}_{A'}$). Indeed (2.8) may be taken as defining Ω^1. The long exact sequence on cohomology which arises from (2.8) gives

$$\begin{array}{ccccccc} \Gamma(\mathbb{P}_1, \mathcal{O}^{A'}(-1)) & \longrightarrow & \Gamma(\mathbb{P}_1, \mathcal{O}) & \longrightarrow & H^1(\mathbb{P}_1, \Omega^1) & \longrightarrow & H^1(\mathbb{P}_1, \mathcal{O}^{A'}(-1)) \\ \| & & & & & & \| \\ 0 \text{ (by (2.7))} & & & & & & 0 \text{ (by (2.7))} \end{array}$$

which shows $H^1(\mathbb{P}_1, \Omega^1)$ is isomorphic to $\Gamma(\mathbb{P}_1, \mathcal{O})$. But now $\Gamma(\mathbb{P}_1, \mathcal{O})$ is canonically isomorphic to $\mathbb{C}$ by evaluation at any point. To mimic this proof for $\mathcal{O}(-2)$ we use instead of (2.8) the exact sequence (isomorphic to (2.8), but not canonically):

$$\mathbb{P}_1 ; \; 0 \longrightarrow \mathcal{O}(-2) \xrightarrow{\pi_{A'}} \mathcal{O}_{A'}(-1) \xrightarrow{\pi_{B'}} \mathcal{O}_{[A'B']} \longrightarrow 0. \tag{2.9}$$

This shows that $H^1(\nu^{-1}(z), \mathcal{O}_{\nu^{-1}(z)}(-2))$ is canonically isomorphic to $\mathcal{O}_{[A'B']}$ and thus $\nu_*^1 \mathcal{O}(-2) \cong \mathcal{O}_{[A'B']} = \mathcal{O}[-1]'$, as required. Actually, we could deduce directly from (2.9), considered as a sequence on $\mathbb{F}$, that $\nu_*^1 \mathcal{O}(-2) \cong \nu_* \mathcal{O}_{[A'B']} \cong \mathcal{O}_{[A'B']}$ just by using the long exact sequence of direct images [45]. This would avoid explicit use of (2.6). To identify $\nu_*^1 \mathcal{O}(-n-2)$ we replace (2.9) with

$$\mathbb{P}_1 ; \; 0 \longrightarrow \mathcal{O}(-n-2) \xrightarrow{\pi_{A'}\cdots\pi_{E'}} \mathcal{O}_{(A'B'\cdots D'E')}(-1) \xrightarrow{\pi_{F'}} \mathcal{O}_{(A'B'\cdots D')[E'F']} \longrightarrow 0 \tag{2.10}$$

to conclude $\nu_*^1 \mathcal{O}(-n-2) = \mathcal{O}_{(A'B'\cdots D')[E'F']} = \mathcal{O}_{(A'B'\ldots D')}[-1]'$. □

In the proof of the above proposition we showed that if U is an open Stein subset of $\mathbb{M}$ then $H^q(U', \mathcal{O}(k))$ is isomorphic to $\Gamma(U, \nu_*^q \mathcal{O}(k))$. More generally we have:

Proposition 2.2. *Let $U \subseteq \mathbb{M}$ be any open set. Then, for all p,*

a) if $n \geqq 0$: $H^p(U', \mathcal{O}(n)) \cong H^p(U, \mathcal{O}^{(A'B'\cdots D')})$
b) $H^p(U', \mathcal{O}(-1)) = 0$
c) if $n \geqq 0$: $H^p(U', \mathcal{O}(-n-2)) \cong H^{p-1}(U, \mathcal{O}_{(A'B'\ldots D')}[-1]')$.

Proof. These statements follow immediately from the Leray spectral sequence [14] and Proposition 2.1 which shows that $E_2^{p,q} = E_\infty^{p,q}$. Alternatively one can argue directly using Čech cohomology and a Stein cover of U. □

From Proposition 2.1 and Lemma 1.2 it follows immediately that

$$
\begin{aligned}
&\nu_* \Omega_\mu^0(n) \cong \nu_* \mathcal{O}(n) \cong \mathcal{O}^{(A'B'\cdots D')} \quad (n \text{ indices}) \quad \text{if } n \geqq 0\\
&\nu_* \Omega_\mu^1(n) \cong \nu_* \mathcal{O}_A(n+1)[-1]' \cong \mathcal{O}_A^{(A'B'\cdots E')}[-1]' \quad (n+1 \text{ indices}) \quad \text{if } n \geqq -1\\
&\nu_* \Omega_\mu^2(n) \cong \nu_* \mathcal{O}(n+2)[-1][-2]' \cong \mathcal{O}^{(A'B'\cdots F')}[-1][-2]' \quad (n+2 \text{ indices})\\
&\mathbb{M}; \qquad\qquad \text{if } n \geqq -2\\
&\nu_*^1 \Omega_\mu^0(-n-2) \cong \nu_*^1 \mathcal{O}(-n-2) \cong \mathcal{O}_{(A'B'\cdots D')}[-1]' \quad (n \text{ indices}) \quad \text{if } n \geqq 0\\
&\nu_*^1 \Omega_\mu^1(-n-2) \cong \nu_*^1 \mathcal{O}_A(-n-1)[-1]' \cong \mathcal{O}_{A(B'\cdots D')}[-2]' \quad (n-1 \text{ indices})\\
&\qquad\qquad \text{if } n \geqq 1\\
&\nu_*^1 \Omega_\mu^2(-n-2) \cong \nu_*^1 \mathcal{O}(-n)[-1][-2]' \cong \mathcal{O}_{(C'\cdots D')}[-1][-3]' \quad (n-2 \text{ indices})\\
&\qquad\qquad \text{if } n \geqq 2 \qquad (2.11)
\end{aligned}
$$

and all other possibilities vanish. The obvious analogue of Proposition 2.2 holds. Our next task is to interpret the effect of the differential operators d_μ on direct images. For ν_* we can argue as follows. Let $f^{A'B'\cdots D'}$ be a local section of $\mathcal{O}^{(A'B'\cdots D')}$. The differential operator

$$\mathbb{F};\ \pi_{E'} \nabla_A^{E'} : \mathcal{O}(n) \longrightarrow \mathcal{O}_A(n+1)[-1]'$$

gives

$$\pi_{A'} \pi_{B'} \ldots \pi_{D'} f^{A'B'\cdots D'} \longmapsto \pi_{A'} \pi_{B'} \ldots \pi_{D'} \pi_{E'} \nabla_A^{E'} f^{A'B'\cdots D'}$$

which induces on $\nu_* \mathcal{O}(n) = \mathcal{O}^{(A'B'\cdots D')}$

$$\nabla_A^{E'} : f^{A'\cdots D'} \longmapsto \nabla_A^{(E'} f^{A'B'\cdots D')}.$$

Hence, Lemma 1.2 and (2.11) give

$$
\begin{array}{ccccc}
\nu_* \Omega_\mu^0(n) & \xrightarrow{d_\mu} & \nu_* \Omega_\mu^1(n) & \xrightarrow{d_\mu} & \nu_* \Omega_\mu^2(n)\\
\wr\| & & \wr\| & & \wr\| \\
\mathbb{M};\ \mathcal{O}^{(A'B'\cdots D')} & \longrightarrow & \mathcal{O}_A^{(A'B'\cdots E')}[-1]' & \longrightarrow & \mathcal{O}^{(A'B'\cdots F')}[-1][-2]'\\
\cup & & \cup & & \cup \\
f^{A'B'\cdots D'} & \longmapsto & \nabla_A^{(E'} f^{A'B'\cdots D')}, g_A^{A'B'\cdots E'} & \longmapsto & \nabla^{A(F'} g_A^{A'B'\cdots E')}
\end{array}
\qquad (2.12)
$$

Notice that this shows automatically that the operator $\nabla_A^{E'}$ from $\mathcal{O}^{(A'B'\cdots D')}$ to $\mathcal{O}_A^{(A'B'\cdots E')}[-1]'$ is well-defined. To compute the effect of d_μ on the first direct images we can use (2.10) and write locally, after choosing coordinates $\pi_{A'}$:

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

$$\begin{array}{ccccccccc}
 & & 0 & \longrightarrow & \mathcal{O}(-n-2) & \xrightarrow{\pi_{A'}\cdots\pi_{E'}} & \mathcal{O}_{(A'B'\cdots D'E')}(-1) & \xrightarrow{\pi^E{}_{F'}\pi^F{}_{G'}} & \mathcal{O}_{(A'B'\cdots D'(E'F'G'))} \longrightarrow 0 \\
 & & \downarrow{\scriptstyle \pi_A{}^{A'}\nabla_{A'}^{A}} & & & & \downarrow{\scriptstyle \nabla_A^A} & & \downarrow{\scriptstyle \nabla_A^A} \\
\mathbb{N}: & 0 & \longrightarrow \mathcal{O}_A(-n-1)[-1]' & \xrightarrow{\pi_{B'}\cdots\pi_{E'}} & \mathcal{O}_{A(B'\cdots D'E')}(-1)[-1]' & \xrightarrow{\pi^E{}_{F'}\pi^F{}_{G'}} & \mathcal{O}_{A(B'\cdots D'(F'G'))}[-1]' & \longrightarrow 0 &
\end{array}$$

Hence

$$\begin{array}{ccccccc}
\nu_*^1\mathcal{O}(-n-2) \simeq \nu_*\mathcal{O}_{(A'B'\cdots D'(E'F'))} & = & \mathcal{O}_{(A'B'\cdots D'(E'F'))} & = & \mathcal{O}_{(A'B'\cdots D')}[-1]' & & \\
\downarrow{\scriptstyle \nu_*^1(\pi_A{}^{A'}\nabla_{A'}^A)} & & \downarrow{\scriptstyle \nu_*(\nabla_A^A)} & & \downarrow{\scriptstyle \nabla_A^A} & & \downarrow{\scriptstyle \nabla_A^A} \\
\nu_*^1\mathcal{O}_A(-n-1)[-1]' \cong \nu_*\mathcal{O}_{A(B'\cdots D'(E'F'))}[-1]' & = & \mathcal{O}_{A(B'\cdots D'(E'F'))}[-1]' & \cong & \mathcal{O}_{A(B'\cdots D')}[-2] &&
\end{array}$$

and so the analogue of (2.12) for ν_*^1 is

$$\begin{array}{ccccc}
\nu_*^1\Omega_\mu^0(-n-2) & \xrightarrow{d_\mu} & \nu_*^1\Omega_\mu^1(-n-2) & \xrightarrow{d_\mu} & \nu_*^1\Omega_\mu^2(-n-2) \\
\| & & \| & & \| \\
\mathbb{M}: \mathcal{O}_{(A'B'\cdots D')}[-1]' & \longrightarrow & \mathcal{O}_{A(B'\cdots D')}[-2]' & \longrightarrow & \mathcal{O}_{(C'\cdots D')}[-1][-3]' \\
\phi_{A'B'\cdots D'} & \longmapsto & \nabla_A^{A'}\phi_{A'B'\cdots D'}, \psi_{AB'C'\cdots D'} & \longmapsto & \nabla^{AB'}\psi_{AB'C'\cdots D'}.
\end{array} \qquad (2.13)$$

We emphasize that the formulae of (2.12) and (2.13) arise quite naturally just from the basic twistor geometry of (1.1). These formulae give rise to massless fields as we shall see. In other words, the massless field equations are implicitly generated by the twistor geometry. This would seem to be a significant simplification.

3. The Topology of the Mapping μ

We want to be able to interpret analytic information on subsets of $\mathbb{P}$ (for instance, certain cohomology groups) in terms of similar objects defined on subsets of $\mathbb{F}$. To do this we need to investigate some further properties of $\mu: \mathbb{F} \to \mathbb{P}$.

Let $\tilde{W}$ be an open subset of $\mathbb{F}$ and $W = \mu(\tilde{W})$. W will also be open. We want to study the behaviour of sheaves and cohomology on W and $\tilde{W}$, determining conditions on $\tilde{W}$ so that we can transform, without loss of information, data on W to data on $\tilde{W}$. If $\mathcal{S}$ is a sheaf on W then we may form the *topological inverse image sheaf* $\mu^{-1}\mathcal{S}$ on $\tilde{W}$. It may be defined by the presheaf $V \mapsto \Gamma(\mu(V), \mathcal{S})$ for V open in $\tilde{W}$ with the obvious restriction mappings, and it is characterized by the requirement that $(\mu^{-1}\mathcal{S})_q \cong \mathcal{S}_{\mu(q)}$ for all $q \in \tilde{W}$. This isomorphism gives rise to a map

$$\mu^*: \Gamma(W, \mathcal{S}) \longrightarrow \Gamma(\tilde{W}, \mu^{-1}\mathcal{S}),$$

defined by $\mu^* f = f \circ \mu$ (where, by abuse of notation, we have identified $(\mu^{-1}\mathcal{S})_q$ with $\mathcal{S}_{\mu(q)}$). The sheaf $\mu^{-1}\mathcal{O}(V)$ introduced in Sect. 1 is an example of this construction. The pull-back μ^* may be extended to cohomology in the following manner (cf. Godement [14]). We may calculate $H^p(W, \mathcal{S})$ by choosing a suitable resolution of $\mathcal{S}$, i.e. an exact sequence

$$W: \ 0 \longrightarrow \mathcal{S} \xrightarrow{\delta} \mathcal{R}^0 \xrightarrow{\delta} \mathcal{R}^1 \xrightarrow{\delta} \mathcal{R}^2 \xrightarrow{\delta} \cdots$$

which we'll abbreviate to

$$W: \ 0 \longrightarrow \mathcal{S} \longrightarrow \mathcal{R}^\cdot, \qquad (3.1)$$

where "suitable" can be taken to mean injective, flabby, fine or soft (often, in practice, a resolution by sheaves of smooth differential forms). Then the abstract deRham theorem (cf. Wells [51]) says that there is a canonical isomorphism

$$H^p(W, \mathscr{S}) = H^p(\Gamma(W, \mathscr{R}^\cdot)) := \frac{\ker \delta : \Gamma(W, \mathscr{R}^p) \longrightarrow \Gamma(W, \mathscr{R}^{p+1})}{\operatorname{im} \delta : \Gamma(W, \mathscr{R}^{p-1}) \longrightarrow \Gamma(W, \mathscr{R}^p)} \tag{3.2}$$

The resolution of $\mathscr{S}$ in (3.1) gives rise to a resolution

$$W;\ 0 \longrightarrow \mu^{-1}\mathscr{S} \longrightarrow \mu^{-1}\mathscr{R}^\cdot$$

of $\mu^{-1}\mathscr{S}$. By writing this as a collection of short exact sequences and looking at the corresponding long exact sequences on cohomology (or by using some universal property of resolutions) it may easily be seen that there is a canonical homomorphism

$$H^p(\Gamma(\tilde{W}, \mu^{-1}\mathscr{R}^\cdot)) \longrightarrow H^p(\tilde{W}, \mu^{-1}\mathscr{S}). \tag{3.3}$$

The mapping μ^* on sections gives rise to a map of complexes

$$\mu^* : \Gamma(W, \mathscr{R}^\cdot) \longrightarrow \Gamma(\tilde{W}, \mu^{-1}\mathscr{R}^\cdot)$$

and hence a map on cohomology which we can combine with (3.2) and (3.3) to define μ^* on $H^p(W, \mathscr{S})$ by means of the commutative diagram

$$\begin{array}{ccc} H^p(W, \mathscr{S}) & \xrightarrow{\mu^*} & H^p(\tilde{W}, \mu^{-1}\mathscr{S}) \\ \| & & \uparrow \\ H^p(\Gamma(W, \mathscr{R}^\cdot)) & \longrightarrow & H^p(\Gamma(\tilde{W}, \mu^{-1}\mathscr{R}^\cdot)). \end{array}$$

It is easy to check that this construction is independent of the choice of resolution, $\mathscr{R}^\cdot$. It is also not difficult to see how to define μ^* naturally in terms of Čech theory.

If $\mu : \tilde{W} \to W$ has connected fibres, it is clear that

$$\Gamma(W, \mathscr{S}) = \Gamma(\tilde{W}, \mu^{-1}\mathscr{S}).$$

The situation is not so simple for the higher cohomology,

$$\mu^* : H^p(W, \mathscr{S}) \longrightarrow H^p(\tilde{W}, \mu^{-1}\mathscr{S}), \quad p \geqq 1. \tag{3.4}$$

As it turns out, the higher order topology of the fibres of μ must enter into the discussion. The basic condition we shall use is that the mapping

$$\mu : \tilde{W} \longrightarrow W$$

have fibres which are connected and have vanishing first Betti number. In other words, for all $Z \in W$, $\mu^{-1}(Z)$ is connected and $H^1(\mu^{-1}(Z), \mathbb{C}) = 0$. We shall refer to such a mapping by saying it is *elementary*.

Remark. If $U \subseteq \mathbb{M}$ and U', U'' are as in Sect. 1 the fibre $\mu^{-1}(Z)$ of $\mu : U' \to U''$ is biholomorphic to $\hat{Z} \cap U$ in $\mathbb{M}$. Here $\hat{Z}$ is the α-plane determined by the (projective) twistor $Z \in \mathbb{P}$ (see Sect. 1). Our principal interest later will be where $W = U''$

and $\tilde{W} = U'$ so this observation may be used to check whether $\mu: \tilde{W} \to W$ is elementary.

Returning to the question of the nature of the mapping (3.4), we consider the special case of $\mathscr{S} = \mathscr{O}(V)$ for some holomorphic vector-bundle V on W. We may define μ^* by using the Dolbeault resolution

$$W;\ 0 \longrightarrow \mathscr{O}(V) \longrightarrow \mathscr{E}^{0,0}(V) \xrightarrow{\bar{\partial}} \mathscr{E}^{0,1}(V) \xrightarrow{\bar{\partial}} \mathscr{E}^{0,2}(V) \xrightarrow{\bar{\partial}} \dots$$

where $\mathscr{E}^{p,q}(V)$ denotes the sheaf of forms of type (p, q) with coefficients in V (cf. [51] or [43]). If V is a smooth (i.e. C^∞) vector-bundle we let $\mathscr{E}(V)$ be the sheaf of smooth sections of V (i.e. $\mathscr{E}(V) = \mathscr{E}^{0,0}(V)$).

Lemma 3.1. *Suppose that $\mu: \tilde{W} \to W$ is elementary, and that V is a smooth vector bundle over W. Then*

$$H^1(\tilde{W}, \mu^{-1}\mathscr{E}(V)) = 0.$$

Proof. By analogy with the holomorphic case of Sect. 1 we may define a sheaf of smooth relative p-forms $\mathscr{E}^p_\mu$ on $\tilde{W}$ and differential operators d_μ so that we obtain a fine resolution

$$\tilde{W};\ 0 \longrightarrow \mu^{-1}\mathscr{E}(V) \longrightarrow \mathscr{E}^0_\mu(V) \xrightarrow{d_\mu} \mathscr{E}^1_\mu(V) \xrightarrow{d_\mu} \mathscr{E}^2_\mu(V) \xrightarrow{d_\mu} \mathscr{E}^3_\mu(V) \xrightarrow{d_\mu} \mathscr{E}^4_\mu(V) \longrightarrow 0 \quad (3.5)$$

(cf. (1.5)). Suppose that $Z \in W$. We want to show that there is a neighbourhood U_Z of Z so that $H^1(\mu^{-1}(U_Z) \cap \tilde{W}, \mu^{-1}\mathscr{E}(V)) = 0$. First, we may choose U_Z so that $\mu^{-1}(U_Z)$ is trivialized:

$$\begin{array}{ccc} \mathbb{F} \supseteq \mu^{-1}(U_Z) & \simeq & U_Z \times \mathbb{P}_2 \\ \downarrow \mu & & \swarrow \\ \mathbb{P} \supseteq & & U_Z \end{array}$$

Next, we choose any $s_0 \in \mathbb{P}_2$ such that $(Z, s_0) \in \tilde{W}$ (where, by abuse of notation we have identified $\mu^{-1}(U_Z)$ and $U_Z \times \mathbb{P}_2$) and shrink U_Z, if necessary, to ensure that $X \in U_Z$ implies $(X, s_0) \in \tilde{W}$. We may also suppose that V is trivial on $\tilde{U}_Z$, where $\tilde{U}_Z$ denotes $\mu^{-1}(U_Z) \cap \tilde{W}$. We have

$$H^1(\tilde{U}_z, \mu^{-1}\mathscr{E}(V)) \cong H^1(\tilde{U}_z, \mu^{-1}\mathscr{E})^r,$$

where r is the rank of V and $\mathscr{E}$ is the sheaf of smooth functions on $\mathbb{P}$. As a special case of (3.5) we see that $H^1(\tilde{U}_z, \mu^{-1}\mathscr{E}) = 0$ if and only if

$$\Gamma(\tilde{U}_z, \mathscr{E}^0_\mu) \xrightarrow{d_\mu} \Gamma(\tilde{U}_z, \mathscr{E}^1_\mu) \xrightarrow{d_\mu} \Gamma(\tilde{U}_z, \mathscr{E}^2_\mu)$$

is exact. So suppose $\omega \in \Gamma(\tilde{U}_z, \mathscr{E}^1_\mu)$ with $d_\mu \omega = 0$. Then ω may be considered as a family of smooth closed 1-forms $\omega(X)$ in the s-variables ($s \in \mathbb{P}_2$) smoothly parametrized by $X \in U_z$. Thus, we may define $f \in \Gamma(\tilde{U}_z, \mathscr{E}^0_\mu)$ by means of the integral formula

$$f(X, s) = \int_{(X, s_0)}^{(X, s)} \omega(X)$$

where the integral is taken over any path in $\mu^{-1}(X)\cap\tilde{W}$ joining (X, s_0) to (X, s). This is possible and independent of the choice of path since $\mu:\tilde{W}\to W$ is elementary. By construction $d_\mu f=\omega$ so $H^1(\tilde{U}_z, \mu^{-1}\mathscr{E})=0$. Hence $H^1(\tilde{U}_z, \mu^{-1}\mathscr{E}(V))=0$ i.e.

$$\Gamma(\tilde{U}_z, \mathscr{E}^0_\mu(V)) \xrightarrow{d_\mu} \Gamma(\tilde{U}_z, \mathscr{E}^1_\mu(V)) \xrightarrow{d_\mu} \Gamma(\tilde{U}_z, \mathscr{E}^2_\mu(V)) \tag{3.6}$$

is exact. Choose a covering of W by sets U_z of the above form and a partition of unity subordinate to this covering. When pulled back to $\tilde{W}$ this partition of unity is constant along the fibres of μ, i.e. is annihilated by d_μ. It may therefore be used to patch solutions to the equation "(3.6) is exact" to obtain the exactness of

$$\Gamma(\tilde{W}, \mathscr{E}^0_\mu(V)) \xrightarrow{d_\mu} \Gamma(\tilde{W}, \mathscr{E}^1_\mu(V)) \xrightarrow{d_\mu} \Gamma(\tilde{W}, \mathscr{E}^2_\mu(V))$$

or, in other words, (by (3.5)) $H^1(\tilde{W}, \mu^{-1}\mathscr{E}(V))=0$, as desired. □

Remark. If $W=U''$, $\tilde{W}=U'$, where $U\subset\mathbb{M}$ is Stein, then the above argument can be used directly on the sequence (1.6) to prove the following theorem in the special case $V=H^n, n\geq -1$. For $n=0$ (i.e. V is the trivial bundle) the proof is essentially given in Ward [47].

Theorem 3.2. *Suppose $\tilde{W}\subseteq\mathbb{F}$ is an open set and let W denote $\mu(\tilde{W})$. If $\mu:\tilde{W}\to W$ is elementary then, for any holomorphic vector-bundle V on W,*

$$\mu^*: H^1(W, \mathscr{O}(V)) \longrightarrow H^1(\tilde{W}, \mu^{-1}\mathscr{O}(V))$$

is a canonical isomorphism.

Proof. The sheaves $\mathscr{E}^{p,q}(V)$ are of the form $\mathscr{E}(\tilde{V})$ for $\tilde{V}=\Lambda^{p,q}T^*(W)\otimes V$, where $\Lambda^{p,q}T^*(W)$ is the smooth bundle whose sections are forms of type (p,q). Thus by Lemma 3.1 $H^1(\tilde{W}, \mu^{-1}\mathscr{E}^{p,q}(V))=0$ for all p,q. Actually we only need $H^1(\tilde{W}, \mu^{-1}\mathscr{E}(V))=0$ to conclude by simple diagram chasing that the canonical map (3.3)

$$H^1(\Gamma(\tilde{W}, \mu^{-1}\mathscr{E}^{0,\cdot}(V)) \longrightarrow H^1(\tilde{W}, \mu^{-1}\mathscr{O}(V))$$

is an isomorphism (it's always injective). Now, $\mu:\tilde{W}\to W$ has connected fibres so

$$\Gamma(\tilde{W}, \mu^{-1}\mathscr{E}^{0,q}(V)) = \Gamma(W, \mathscr{E}^{0,q}(V)).$$

Hence, μ^* is the composition of three isomorphisms

$$H^1(W, \mathscr{O}(V)) \xrightarrow{\cong} H^1(\Gamma(W, \mathscr{E}^{0,\cdot}(V)))$$
$$\xrightarrow{\cong} H^1(\Gamma(\tilde{W}, \mu^{-1}\mathscr{E}^{0,\cdot}(V))) \xrightarrow{\cong} H^1(\tilde{W}, \mu^{-1}\mathscr{O}(V))$$

and the result follows. □

The above theorem indicates how the topology of the fibres of μ affects the properties of μ^*. So far we have:

A. $\mu^*:\Gamma(W, \mathscr{O}(V)) \longrightarrow \Gamma(W, \mu^{-1}\mathscr{O}(V))$ is always injective.

B. If $\mu:\tilde{W} \longrightarrow W$ has connected fibres then
$\mu^*:\Gamma(W, \mathscr{O}(V)) \longrightarrow \Gamma(\tilde{W}, \mu^{-1}\mathscr{O}(V))$ is an isomorphism and
$\mu^*:H^1(W, \mathscr{O}(V)) \longrightarrow H^1(\tilde{W}, \mu^{-1}\mathscr{O}(V))$ is injective.

C. If, in addition, the first Betti number of each fibre of $\mu: \tilde{W} \longrightarrow W$ vanishes (in other words, $\mu: \tilde{W} \longrightarrow W$ is elementary), then $\mu^*: H^1(W, \mathcal{O}(V)) \longrightarrow H^1(\tilde{W}, \mu^{-1}\mathcal{O}(V))$ is an isomorphism and $\mu^*: H^2(W, \mathcal{O}(V)) \longrightarrow H^2(\tilde{W}, \mu^{-1}\mathcal{O}(V))$ is injective (this last assertion following easily from Lemma 3.1 by a similar argument to the proof of Theorem 3.2). We expect this to be part of a general pattern but a proof of the general case is lacking*. It is, however, fairly straightforward to show that if all of the fibres of $\mu: \tilde{W} \to W$ are convex, then $\tilde{W}$ is homeomorphic to a product $W \times$ fibre, and one can deduce via a higher degree version of Lemma 3.1 that

$$\mu^*: H^p(W, \mathcal{O}(V)) \xrightarrow{\cong} H^p(\tilde{W}, \mu^{-1}\mathcal{O}(V))$$

for all p. An alternative proof of this in the case $W = \mathbb{P}^+$, $\tilde{W} = \mathbb{F}^+$ is given in Eastwood [9]. For the applications in this paper, however, we shall be concerned primarily with the action of μ^* on H^1 so Theorem 3.2 will suffice.

4. The General $\mathcal{P}$-Transform

The aim of this section is to finish constructing a general analytical machine which transforms elements of analytic cohomology groups $H^p(U'', \mathcal{O}(V))$ into solutions of certain differential equations on U, an open subset of $\mathbb{M}$. The complete process will be called the *$\mathcal{P}$-transform*[9]. The pieces of this transform already constructed include the isomorphism.

$$\mu^*: H^1(U'', \mathcal{O}(V)) \xrightarrow{\cong} H^1(U', \mu^{-1}\mathcal{O}(V)) \tag{4.1}$$

when $\mu: U' \to U''$ is elementary (Sect. 3), and the canonical direct image isomorphisms (for instance):

$$\begin{array}{lll} \text{a)} & H^1(U', \Omega^0_\mu(-n-2)) & \xrightarrow{d_\mu} H^1(U', \Omega^1_\mu(-n-2)) \\ & \| & \| \\ \text{b)} & \Gamma(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') & \xrightarrow{\nabla^A_{\dot{A}}} \Gamma(U, \mathcal{O}_{A(B'\cdots D')}[-2]') \end{array} \tag{4.2}$$

discussed in Sect. 2. What is missing in this case is a relation between $H^1(U', \mu^{-1}\mathcal{O}(-n-2))$ and $H^1(U', \Omega^p_\mu(-n-2))$. This will be a special case of what we will establish in this section (Theorem 4.1). If $n \geqq 1$, it's not difficult to show that

$$H^1(U', \mu^{-1}\mathcal{O}(-n-2) \cong \ker d_\mu \text{ in (4.2a)} \tag{4.3}$$

---

* *Note added in proof.* The general case has recently been proved by N. P. Buchdahl

9 The transform we wish to describe has been referred to elsewhere (e.g. in Wells [53]) as the *Penrose transform*. It has also been referred to (e.g. in Rawnsley [42]) as the *twistor transform*. We prefer to avoid the term twistor transform in this context since it is already in common use as a name for the isomorphism

$$\mathcal{T}: H^1(\mathbb{P}^+, \mathcal{O}(-n-2) \xrightarrow{\cong} H^1(\mathbb{P}^{*-}, \mathcal{O}(n-2))$$

(cf. Penrose-MacCallum [32], where this is expressed in pre-cohomological language)

and thus we have the integral transform (assuming $\mu: U' \to U''$ is elementary)

$$\mathscr{P}: H^1(U'', \mathscr{O}(-n-2)) \xrightarrow{\cong} \ker \nabla_A^{A'} \text{ in (4.2b)} \tag{4.4}$$

The details are in Sect. 5 but this is the basic idea. The general form of (4.3) will be a *spectral sequence* (cf. Eastwood [9], in case $U = \mathbb{M}^+$) linking $H^r(U', \mu^{-1}\mathscr{O}(V))$ and $H^q(U', \Omega_\mu^p(V))$. Depending on the geometry and the bundles involved the spectral sequence will give the right relationships between these cohomology groups (e.g. (4.3) above). In the remaining sections of this paper we shall work out various special cases of the $\mathscr{P}$-transform. In this section we will only formulate the general method.

In general terms the transform works as follows. An element ω in $H^p(U'', \mathscr{O}(V))$ is first pulled-back to $\mu^*\omega$ in $H^p(U', \mu^{-1}\mathscr{O}(V))$. We then restrict the cohomology class $\mu^*\omega$ to a fibre $\nu^{-1}(z)$ for $z \in U$, giving a "value" in the stalk of an appropriate direct-image sheaf. As z varies this is interpreted as a spinor-field. Thus,

$$(\mathscr{P}\omega)(z) = \omega \text{ "evaluated" on } L_z.$$

Indeed, the original integral formulae are of this type without explicit reference to the intermediate space $U' \subseteq \mathbb{F}$. However, it turns out that to break the transform into two steps via $\mathbb{F}$ is a good way of being able the analyze the analytical behaviour of $\mathscr{P}$.

We now have the principal result of this section:

Theorem 4.1. *Let X be an open subset of $\mathbb{F}$. Then there is a spectral sequence*

$$E_1^{p,q} = H^q(X, \Omega_\mu^p(V)) \Longrightarrow H^{p+q}(X, \mu^{-1}\mathscr{O}(V))$$

where the differentials $d_1: E_1^{p,q} \to E_1^{p+1,q}$ are induced by $d_\mu: \Omega_\mu^p(V) \to \Omega_\mu^{p+1}(V)$,

Proof. The exact sequence (1.5):

$$\tilde{W};\ 0 \longrightarrow \mu^{-1}\mathscr{O}(V) \longrightarrow \Omega_\mu^0(V) \longrightarrow \Omega_\mu^1(V) \longrightarrow \Omega_\mu^2(V) \longrightarrow 0$$

provides a resolution of $\mu^{-1}\mathscr{O}(V)$. The spectral sequence called for in the theorem therefore follows a standard construction (for a "differential sheaf" in a special case cf. Bredon [5]). □

Remark. One way of constructing the spectral sequence in Theorem 4.1 is to write (1.5) as a pair of short exact sequences, consider the resulting long exact sequences on cohomology, and do some elementary diagram chasing. The same diagram chasing can, of course, be used to prove directly anything which can be deduced from the spectral sequence. The spectral sequence terminology does produce, however, a more illuminating and unifying way of viewing the results of the next few sections, as well as a means of keeping track of sometimes complicated information.

5. The Twistor Description of Right-handed Fields

Let $\mathscr{O}_{(A'B'\cdots D')}[-1]'$ be the spinor sheaf on $\mathbb{M}$ with $n(\geq 1)$ indices as defined in Sect. 2. Let

$$\mathbb{M};\ \mathscr{Z}_n' := \ker \nabla_A^{A'}: \mathscr{O}_{(A'B'\cdots D')}[-1]' \longrightarrow \mathscr{O}_{A(B'\cdots D')}[-2]'. \tag{5.1}$$

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

We shall call $\mathscr{Z}'_n$ the sheaf of *holomorphic right-handed massless free fields of helicity* $n/2$. The sections of $\mathscr{Z}'_n$ over an open subset U contained in $\mathbb{M}$,

$$\Gamma(U, \mathscr{Z}'_n) = \{\phi_{A'B'\cdots D'} \in \Gamma(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') : \nabla^{A'}_{A}\phi_{A'B'\cdots D'} = 0\}$$

are the *holomorphic massless fields on* U *of helicity* $n/2$ (Dirac [6], Fierz [12]). Note that we do not (as is sometimes done) reserve the term "right-handed" (or equivalently "positive helicity") for the case $U = \mathbb{M}^+$. For helicity $= \frac{1}{2}$ these equations are the Dirac–Weyl equations of the anti-neutrino, for helicity $= 1$ we have Maxwell's equations for a self-dual field (a right-handed photon), and for helicity $= 2$ the self-dual linearized Einstein equations. More precisely the sections of $\mathscr{Z}'_n$ are the holomorphic right-handed massless free fields of helicity $n/2$ with *primed conformal weight*. On $\mathbb{M}^I$ these fields coincide with the classical massless fields first introduced by Dirac. We could also envisage fields with unprimed conformal weight, i.e. we could replace (5.1)′ by

$$\mathbb{M};\ \mathscr{Z}_n := \ker \nabla^{A'}_{A} : \mathcal{O}_{(A'B'\cdots D')}[-1] \longrightarrow \mathcal{O}_{A(B'\cdots D')}[-1][-1]' \qquad (5.1)$$

which also agrees with Dirac's definition for $\mathbb{M}^I$. The only distinction here arises when we consider non-unimodular $GL(\mathbb{T})$ transformations, under which $\mathscr{Z}_n$ and $\mathscr{Z}'_n$ are not *canonically* isomorphic. The physical significance of this is yet unclear, but it turns out that the $\mathscr{P}$-transform acting on $\mathbb{P}$ produce naturally sections of $\mathscr{Z}'_n$. The distinction between $\mathscr{Z}'_n$ and $\mathscr{Z}_n$ would disappear (as would that between primed and unprimed weights generally) if we had been concerned with the action of $SL(\mathbb{T})$ rather than $GL(\mathbb{T})$.

We now have the following theorem which represents right-handed holomorphic massless fields on $\mathbb{M}$ in terms of holomorphic data on $\mathbb{P}$.

Theorem 5.1. *For* U *open in* $\mathbb{M}$ *and* $n \geq 1$, *there is a canonical linear transformation*

$$\mathscr{P} : H^1(U'', \mathcal{O}(-n-2)) \longrightarrow \Gamma(U, \mathscr{Z}'_n).$$

If $\mu : U' \to U''$ *has connected fibres, then* $\mathscr{P}$ *is injective and if* $\mu : U' \to U''$ *is elementary, then* $\mathscr{P}$ *is an isomorphism.*

Before we proceed with the proof of this theorem we recall that $\mu : U' \to U''$ has connected fibres if and only if every α-plane in $\mathbb{M}$ intersects U in a connected set and that $\mu : U' \to U''$ is elementary if and only if all these intersections are connected and have vanishing first Betti number.

Proof of Theorem 5.1. We already know from Sect. 3 that

$$\mu^* : H^1(U'', \mathcal{O}(-n-2)) \longrightarrow H^1(U', \mu^{-1}\mathcal{O}(-n-2))$$

is injective or bijective according to the hypotheses given in the theorem. Thus it suffices to show that

$$H^1(U', \mu^{-1}\mathcal{O}(-n-2)) \cong \Gamma(U, \mathscr{Z}'_n)$$

in a natural way. We do this by using the spectral sequence of Theorem 4.1. By the formulae (2.11) (and the analogue of Proposition 2.2) we compute:

$$E_1^{p,0} = H^0(U', \Omega^p_\mu(-n-2)) = 0 \quad \text{for all } p$$

$$E_1^{0,1} = H^1(U', \Omega_\mu^0(-n-2)) \cong \Gamma(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') \quad (n \text{ indices})$$
$$E_1^{1,1} = H^1(U', \Omega_\mu^1(-n-2)) \cong \Gamma(U, \mathcal{O}_{A(B'\cdots D')}[-2]')$$
etc.

Thus the E_1-level of the spectral sequence of Theorem 4.1 is isomorphic to

$$\begin{array}{ccccccc} H^1(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') & \longrightarrow & \cdots & & & & \\ \uparrow q & & & & & & \\ \Gamma(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') & \xrightarrow{\nabla_A^{A'}} & \Gamma(U, \mathcal{O}_{A(B'\cdots D')}[-2]') & \xrightarrow{\nabla^{AB'}} & \Gamma(U, \mathcal{O}_{(C'\cdots D')}[-1][-3]') & \longrightarrow & \cdot \\ \uparrow & & \uparrow & & \uparrow & & \\ 0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & p \end{array} \tag{5.2}$$

Thus $E_2^{0,1} \cong \ker \nabla_A^{A'} : \Gamma(U, \mathcal{O}_{(A'B'\cdots D')}[-1]') \longrightarrow \Gamma(U, \mathcal{O}_{A(B'\cdots D')}[-2]')$
$= \Gamma(U, \mathscr{Z}_n')$.

In a general spectral sequence argument we would now have to consider $d_2 : E_2^{0,1} \to E_2^{2,0}$ but $E_2^{1,0} = E_1^{1,0} = 0$ so this is just the zero map. Hence $E_2^{0,1} = E_\infty^{0,1}$. Also $E_1^{1,0} = 0$ so $E_\infty^{1,0} = 0$. In general the spectral sequence of Theorem 4.1 gives an exact sequence

$$0 \longrightarrow E_\infty^{1,0} \longrightarrow H^1(U', \mu^{-1}\mathcal{O}(V)) \longrightarrow E_\infty^{0,1} \longrightarrow 0.$$

In our case, however, $E_\infty^{1,0} = 0$ so we can conclude that

$$H^1(U', \mu^{-1}\mathcal{O}(-n-2)) \cong E_\infty^{0,1} \cong E_2^{0,1} \cong \Gamma(U, \mathscr{Z}_n')$$

as required.

The group $GL(\mathbb{T})$ of linear automorphisms of $\mathbb{T}$ (isomorphic to $GL(4, \mathbb{C})$) clearly acts on $\mathbb{P}$ and $\mathbb{M}$ in such a way that the twistor correspondence (1.1) is preserved. Thus, if $\varphi \in GL(\mathbb{T})$ and U is an open subset of $\mathbb{M}$ then $\varphi(U)'' = \varphi(U'')$. φ also preserves $\mathcal{O}(-n-2)$ and $\mathscr{Z}_n'$. Hence it makes sense to ask that the diagram

$$\begin{array}{ccc} \mathscr{P} : H^1(U'', \mathcal{O}(-n-2)) & \longrightarrow & \Gamma(U, \mathscr{Z}_n') \\ \uparrow \varphi & & \uparrow \varphi \\ \mathscr{P} : H^1(\varphi(U''), \mathcal{O}(-n-2)) & \longrightarrow & \Gamma(\varphi(U), \mathscr{Z}_n') \end{array} \tag{5.3}$$

be commutative. This is what it means for $\mathscr{P}$ to be canonical. It is, moreover, clear that (5.3) is commutative since the proof of Theorem 5.1 is free from any choice of basis for $\mathbb{T}$. □

We carried out the argument in the above proof in rather too much detail. For anyone familiar with the language of spectral sequences a glance at (5.2) would suffice. For other examples and applications of spectral sequences see Bredon [5] or Godement [14]. As remarked earlier, the same result easily follows directly from (1.6) and (2.11) without the use of this seemingly cumbersome machinery.

Massless fields are often described by means of a Fourier integral formula (Fierz [12]). It is then possible, in terms of this formula, to define what is meant by a *positive-frequency* (or equivalently, *positive-energy*) field. Wave-functions occurring in nature are positive-frequency. A real-analytic positive-frequency field on real Minkowski space automatically extends to a holomorphic field on

an open neighbourhood of $\overline{\mathbb{M}^+}$ (often taken as the definition of positive-frequency). If $\mathscr{P}$ were an isomorphism for a suitable collection of such neighbourhoods then we'd be able to take a direct limit over all such neighbourhoods to obtain an isomorphism

$$\mathscr{P}: H^1(\overline{\mathbb{P}^+}, \mathcal{O}(-n-2)) \xrightarrow{\cong} \{\text{real-analytic positive-frequency helicity } n/2 \text{ massless fields on real Minkowski-space}\}$$

In fact, $\mathscr{P}$ is an isomorphism for suitable arbitrarily small open neighbourhoods of $\overline{\mathbb{M}^+}$. The proof is essentially the same as for $\mathbb{M}^+$ itself (see Corollary 5.2, below), the usual choice for U (in Penrose [37] and Wells [52] for example). There is another reason why $\mathbb{M}^+$ and $\mathbb{M}^-$ are of particular interest, namely, these are just the type of set required to define hyperfunctions on real Minkowski-space. Therefore it is reasonable to expect that the $\mathscr{P}$-transform can be extended to include hyperfunction fields and this is indeed possible (see Wells [53] and [54]). The result for $\mathbb{M}^+$ is worth stating separately.

Corollary 5.2. *For $n \geq 1$,*

$$\mathscr{P}: H^1(\mathbb{P}^+, \mathcal{O}(-n-2)) \longrightarrow \Gamma(\mathbb{M}^+, \mathscr{Z}'_n)$$

is a canonical isomorphism.

Proof. $\mathbb{M}^+$ is a convex subset of $\mathbb{M}^I \cong \mathbb{C}^4$. Thus, every α-plane intersects $\mathbb{M}^+$ in a convex and, hence, simply-connected piece. □

It is also clear that Theorem 5.1 describes every right-handed field locally since each point of $\mathbb{M}$ has convex neighbourhoods to which the theorem applies.

6. The Wave Equation

The classical scalar *wave equation* in affine Minkowski space $\mathbb{M}^I$ is

$$\Box\phi = 0, \tag{6.1}$$

where ϕ is a scalar field (a function) and $\Box := \nabla^a\nabla_a = \nabla^{AA'}\nabla_{AA'}$. Every component of a massless field satisfies the wave equation and indeed, (6.1) is also called the *helicity zero massless field equation*. Globally a massless field has to have conformal weight -1. In other words ϕ is to be regarded as a section of $\mathscr{Z}_0$ or $\mathscr{Z}'_0$ where

$$\mathscr{Z}_0 := \ker \Box : \mathcal{O}[-1] \longrightarrow \mathcal{O}[-2][-1]'$$
$$\mathscr{Z}'_0 := \ker \Box : \mathcal{O}[-1]' \longrightarrow \mathcal{O}[-1][-2]'. \tag{6.2}$$

We shall see that $\Box$ arises naturally from the twistor geometry. It is interesting to see how a second order operator appears whereas $\nabla^{A'}_{A}$ in Sect. 5 was first order. The spectral sequence shows how this transition occurs.

Theorem 6.1. *For U open in $\mathbb{M}$ there is a canonical linear transformation*

$$\mathscr{P}: H^1(U'', \mathcal{O}(-2)) \longrightarrow \Gamma(U, \mathscr{Z}'_0).$$

If $\mu: U' \to U''$ has connected fibres, then $\mathscr{P}$ is injective and if $\mu: U' \to U''$ is elementary, then $\mathscr{P}$ is an isomorphism.

Proof. As in the proof of Theorem 5.1 it suffices to compute $H^1(U', \mu^{-1}\mathcal{O}(-2))$ by means of Theorem 4.1. Using (2.11) it follows that

$$E_1^{1,q} = H^q(U', \Omega^1_\mu(-2)) = 0 \quad \text{for all } q.$$

Thus $E_2^{p,q} \cong E_1^{p,q}$ and (2.11) shows that the E_2-level is given by

$$\begin{array}{ccccc}
\uparrow q & & & & \\
H^1(U, \mathcal{O}[-1]') & \searrow & 0 & & H^2(U, \mathcal{O}[-1][-2]') \\
\downarrow & & & \searrow & \\
\Gamma(U, \mathcal{O}[-1]') & \xrightarrow{0D} & & & H^1(U, \mathcal{O}[-1][-2]') \\
\downarrow & & & \searrow & \\
0 & \longrightarrow & 0 & \longrightarrow & \Gamma(U, \mathcal{O}[-1][-2]') \longrightarrow p
\end{array}$$

Thus $H^1(U', \mu^{-1}\mathcal{O}(-2)) \cong \ker D : \Gamma(U, \mathcal{O}[-1]') \to \Gamma(U, \mathcal{O}[-1][-2]')$. Since the differentials of the first level of this spectral sequence all vanish and this is the second level, it follows that D is a second order differential operator. To identify D we proceed as follows. Recall that to identify $H^1(U', \mathcal{O}(-2))$ with $\Gamma(U, \mathcal{O}[-1]')$ we use the sequence (cf. (2.9))

$$\mathbb{F};\ 0 \longrightarrow \mathcal{O}(-2) \xrightarrow{\pi_{A'}} \mathcal{O}_{A'}(-1) \xrightarrow{\pi^A} \mathcal{O}[-1]' \longrightarrow 0$$

and the fact that $\Gamma(U', \mathcal{O}[-1]') \cong \Gamma(U, \mathcal{O}[-1]')$. The map of complexes

$$\begin{array}{ccccc}
\mathcal{O}(-2) & \xrightarrow{\pi_{A'}} & \mathcal{O}_{A'}(-1) & \xrightarrow{\pi^A} & \mathcal{O}[-1]' \\
\downarrow & & \downarrow{\scriptstyle \nabla^{A'}_A} & & \downarrow{\scriptstyle \frac{1}{2}\Box} \\
\mathcal{O}(-2) & \xrightarrow{\pi_A \nabla^A_{A'}} & \mathcal{O}_A(-1)[-1]' & \xrightarrow{\pi_A \nabla^{AA'}} & \mathcal{O}[-1][-2]'
\end{array}$$

gives rise to a map of spectral sequences which on the E_2-level is

$$\begin{array}{ccccc}
\uparrow q & & & & \\
H^1(U', \mathcal{O}(-2)) & \searrow & \cong & & \\
\downarrow & & & \searrow & \\
0 & \longrightarrow & 0 & \longrightarrow & \Gamma(U', \mathcal{O}[-1]') \longrightarrow p \\
q \uparrow \| & & & & \\
H^1(U', \mathcal{O}(-2)) & \searrow & D & & \downarrow{\scriptstyle \frac{1}{2}\Box} \\
\downarrow & & & \searrow & \\
0 & \longrightarrow & 0 & \longrightarrow & \Gamma(U', \mathcal{O}[-1][-2]') \longrightarrow p
\end{array}$$

Hence, $D = \frac{1}{2}\Box$ and the result follows. □

7. Left-handed Fields via Potentials

A classical *holomorphic left-handed massless free field* on an open subset U of $\mathbb{M}$ is a holomorphic spinor field

$$\psi_{AB\cdots D} \in \Gamma(U, \mathcal{O}_{(AB\cdots D)}[-1]') \quad \text{satisfying} \tag{7.1}$$
$$\nabla^{DD'}\psi_{AB\cdots D} = 0,$$

and the field is said to be of helicity $-n/2$ if there are n indices. For helicity -1 these are the spinor form of Maxwell's equations for an anti-self-dual field. We shall denote the sheaf of such by $\mathscr{Z}'_{-n}$. There is also the isomorphic sheaf $\mathscr{Z}_{-n}$ defined as the kernel of $\nabla^{DD'}: \mathcal{O}_{(AB\cdots D)}[-1] \to \mathcal{O}_{(AB\cdots C)}[-2][-1]'$.

One way of representing these fields is via analytic cohomology on portions of dual projective twistor-space $\mathbb{P}^*$. This would follow exactly the construction of Sect. 5 and would yield sections of $\mathscr{Z}_{-n}$. Left-handed fields (sections of $\mathscr{Z}'_{-n}$) may, however, be represented on portions of $\mathbb{P}$ itself. This representation may be accomplished via the contour integral approach (cf. Penrose [34]) or by "power series expansion" of cohomology classes about the lines L_z, for $z \in \mathbb{M}$ (cf. Wells [53] and the discussion in the next section). The method which comes out of Theorem 4.1, however, is to use potentials.

By a *potential* for a field $\psi_{AB\cdots D} \in \Gamma(U, \mathscr{Z}'_{-n})$ we mean a spinor field $\phi_A^{B'\cdots D'} \in \Gamma(U, \mathcal{O}_A^{(B'\cdots D')}[-1]')$ such that

$$\begin{cases} \nabla^{A(A'}\phi_A^{B'\cdots D')} = 0 \\ \nabla_{D'(D}\cdots\nabla_{B'B}\phi_{A)}^{B'\cdots D'} = \psi_{AB\cdots D}. \end{cases} \tag{7.2}$$

(The symmetrization on the left-hand side refers only to the unprimed indices) It is straightforward but tedious to check (cf. (7.6) and (7.7)) that if $\psi_{AB\cdots D}$ is given by (7.2) then the field equations (7.1) are automatically satisfied. There is clearly some "gauge freedom" in choice of potential for it is easy to see that if $\phi_A^{B'\cdots D'}$ is a potential for a field $\psi_{AB\cdots D}$, then so is

$$\phi_A^{B'\cdots D'} + \nabla_A^{(B'}\gamma^{C'\cdots D')} \tag{7.3}$$

for any spinor field $\gamma^{C'\cdots D'} \in \Gamma(U, \mathcal{O}^{C'\cdots D'})$. There are other related notions of potentials. A useful concept, for example, is that of a Hertz potential (cf. Penrose [27]).

Locally it is always possible to find a potential for a given massless field and (7.3) is the only freedom allowed in such a choice. This follows from constructions in Penrose [37] and Ward [49], which involve explicit formulae for representations of the fields and the potentials, respectively, as transforms of the cohomology classes.

According to our definition, a potential for a neutrino field (helicity $-1/2$) is the field itself, so the simplest case is the Maxwell field:

$$\nabla^{BB'}\psi_{AB} = 0. \tag{7.4}$$

If we regard ψ_{AB} as an anti-self-dual 2-form $F_{ab} = \psi_{AB}\varepsilon_{A'B'}$, then (7.4) becomes $dF = 0$, the desired potential is a 1-form ω such that $d\omega = F$, and the gauge freedom is simply that ω may be replaced by $\omega + d\gamma$ for any holomorphic function γ. Hence, the local equivalence of the field potential modulo gauge follows from the exactness of the deRham sequence

$$\mathbb{M};\ \Omega^0 \longrightarrow \Omega^1 \longrightarrow \Omega^2 \longrightarrow \Omega^3.$$

In terms of spinors this portion of the deRham sequence, up to constant factors, is (cf. (2.3))

$$\mathbb{M},\ \mathcal{O} \xrightarrow{\nabla_A^{B'}} \mathcal{O}_A^{B'}[-1]' \xrightarrow[\nabla_{BB'}]{\nabla^{AA'}} \begin{matrix} \mathcal{O}^{(A'B')}[-1][-2]' \\ \oplus \\ \mathcal{O}_{(AB)}[-1]' \end{matrix} \xrightarrow[\nabla^{AA}]{-\nabla_{BB'}} \mathcal{O}_B^{A'}[-1][-2]'. \tag{7.5}$$

Thus, one way of seeing that a left-handed Maxwell field is locally equivalent to a potential modulo gauge, is by studying the following commutative diagram (cf. Eastwood [9]).

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & & & \\
 & & \uparrow & & \uparrow & & & & \\
 & & \mathcal{O} & \xrightarrow{\nabla_A^{B'}} & \mathcal{O}_A^{B'}[-1]' & \xrightarrow{\nabla^{AA'}} & \mathcal{O}^{(A'B')}[-1][-2]' & & \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\mathbb{M}, & \mathcal{O} & \xrightarrow{\nabla_A} & \mathcal{O}_A^{B'}[-1]' & \xrightarrow[\nabla_{BB'}]{\nabla^{AA'}} & \mathcal{O}^{(A'B')}[-1][-2]' \oplus \mathcal{O}_{(AB)}[-1]' & \xrightarrow[\nabla^{AA'}]{\nabla_{BB'}} & \mathcal{O}_B^{A'}[-1][-2]'. & (7.6) \\
 & & & & \uparrow & & \uparrow & & \uparrow \\
 & & & & 0 & \longrightarrow & \mathcal{O}_{(AB)}[-1]' & \xrightarrow{\nabla^{BB'}} & \mathcal{O}_A^{B'}[-1][-2]' \\
 & & & & & & \uparrow & & \uparrow \\
 & & & & & & 0 & & 0
\end{array}$$

The point is that all the columns are exact and the middle row is (7.5) and hence exact. Therefore, simple diagram chasing shows that the cohomology of the first row (potential modulo gauge) is isomorphic to the cohomology of the last row (fields). The reason for rewriting the argument this way is that it generalizes locally to apply to all helicities. The precise details are rather messy and will be omitted. The diagram for helicity $-n/2$ contains $n-1$ intermediate rows between the potential and the field descriptions, the jth intermediate row being the deRham sequence shunted $j-1$ steps to the right and tensored by $\mathcal{O}_{(A_1\cdots A_{n-j-1})}^{(A_1\cdots A_{j-1})}$.

We now consider what happens globally in the helicity -1 case. To investigate the global question of whether fields are equivalent to potentials modulo gauge on a given open subset U of $\mathbb{M}$ we take sections of (7.6) over U. For example, if we were dealing with smooth forms, then sections of (7.5) over U would yield an exact sequence if and only if $H^1(U,\mathbb{C}) = H^2(U,\mathbb{C}) = 0$. Indeed, (7.6) would give rise to an exact sequence

$$H^1(U,\mathbb{C}) \longrightarrow \frac{\{\text{Potentials}\}}{\{\text{Gauge}\}} \longrightarrow \left\{\begin{array}{c}\text{Left-handed Maxwell}\\ \text{fields on } U\end{array}\right\} \longrightarrow H^2(U,\mathbb{C}) \tag{7.7}$$

and so we would conclude that if $H^1(U,\mathbb{C}) = H^2(U,\mathbb{C}) = 0$, then we'd have an isomorphism on U:

$$\frac{\{\text{Potentials}\}}{\{\text{Gauge}\}} \xrightarrow{\cong} \left\{\begin{array}{c}\text{Left-handed}\\ \text{Maxwell fields}\end{array}\right\}.$$

In the case of holomorphic forms, however, it is easy to check that in addition we need

$$\nabla_A^{B'} : H^1(U,\mathcal{O}) \longrightarrow H^1(U,\mathcal{O}_A^{B'}[-1]') \tag{7.8}$$

to be injective. If $U \subseteq \mathbb{M}^I$ is convex then $H^1(U,\mathbb{C}) = H^2(U,\mathbb{C}) = 0$ and U is Stein so that $H^1(U,\mathcal{O}) = 0$. Hence, a potential description is always valid for such a U and, in particular, for $U = \mathbb{M}^+$. The investigation of the general helicity case is similar and will be omitted from this paper. The main result of this section is

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

Theorem 7.1. *Suppose that* $n \geqq 1$, U *is open in* $\mathbb{M}$, *and*

$$\nabla_A^{B'} : H^1(U, \mathcal{O}^{(\overbrace{C\cdots D}^{n-2})}) \longrightarrow H^1(U, \mathcal{O}_A^{(B'C'\cdots D')}[-1]') \tag{7.9}$$

is injective, where this condition is taken to be vacuous if $n = 1$. *Then there is a canonical linear transformation*

$$\mathcal{P} : H^1(U'', \mathcal{O}(n-2)) \longrightarrow \left\{ \begin{array}{c} \text{Potential for helicity } -n/2 \\ \text{holomorphic massless fields on } U \end{array} \right\} \Big/ \left\{ \begin{array}{c} \text{Gauge} \\ \text{Freedom} \end{array} \right\} \tag{7.10}$$

If $\mu : U' \to U''$ *has connected fibres, then* $\mathcal{P}$ *is injective and if* $\mu : U' \to U''$ *is elementary, then* $\mathcal{P}$ *is an isomorphism*

Proof. As in the proofs of Theorems 5.1 and 6.1, $\mathcal{P}$ is constructed by first mapping $\mu^* : H^1(U'', \mathcal{O}(n-2)) \to H^1(U', \mu^{-1}\mathcal{O}(n-2))$, and then using the spectral sequence of Theorem 4.1 to compute $H^1(U', \mu^{-1}\mathcal{O}(n-2))$. Thus it suffices to show that there is a natural isomorphism

$$H^1(U', \mu^{-1}\mathcal{O}(n-2)) \cong \{\text{Potentials}\}/\{\text{Gauge}\}.$$

Using the isomorphisms (2.11) we compute the E_1-level of the spectral sequence:

$$\begin{array}{ccccc} {\scriptstyle q}\uparrow & & & & \\ H^1(U, \mathcal{O}^{(\overbrace{C\cdots D}^{n-2})}) & \xrightarrow{\nabla_A^{B'}} & H^1(U, \mathcal{O}_A^{(B'\cdots D')})[-1]') & \longrightarrow & \cdots \\ \downarrow & & & & \\ \Gamma(U, \mathcal{O}^{(C'\cdots D')}) & \xrightarrow{\nabla_A^{B'}} & \Gamma(U, \mathcal{O}^{(B'\cdots D')})[-1]') & \xrightarrow{\nabla^{AA'}} & \Gamma(U, \mathcal{O}^{(A'B'\cdots D')}[-1][-2]') \xrightarrow[p]{} \cdots \end{array}$$

The injectivity of (7.9) is therefore precisely that $E_2^{0,1} = 0$. Moreover, by definition

$$E_2^{1,0} = \{\text{Potentials}\}/\{\text{Gauge}\}.$$

$E_\infty^{0,1} = E_2^{0,1} = 0$ and $E_\infty^{1,0} \cong E_2^{1,0}$ so the exact sequence

$$0 \longrightarrow E_\infty^{1,0} \longrightarrow H^1(U', \mu^{-1}\mathcal{O}(n-2)) \longrightarrow E_\infty^{0,1} \longrightarrow 0$$

gives the desired result. □

The hypothesis in Theorem 7.1 that (7.9) be injective may seem a little strange. It is a very weak condition and, in particular, Theorem 7.1 applies to $\mathbb{M}^+$ and applies locally in the sense that any point has arbitrarily small neighbourhoods for which the result holds. Nevertheless, we observe that in case $n = 2$ (i.e. left-handed photons) (7.8) and (7.9) coincide. In other words, the extra hypothesis required in the theorem is precisely the additional analytic condition needed, in addition to topological conditions, to ensure the existence of potentials. This seems to indicate that if we drop this condition then it still may be possible to produce the field, but not necessarily a potential (compare also Penrose [37]). In the remainder of this section we shall discuss only the case $n = 2$.

We recall that $\mathcal{P}$ of Theorem 7.1 is constructed by first applying $\mu^* : H^1(U'', \mathcal{O}) \to H^1(U', \mu^{-1}\mathcal{O})$, and then using the spectral sequence of Theorem 4.1 to interpret

$H^1(U', \mu^{-1}\mathcal{O})$ in terms of data on U. If we define a complex of sheaves $\mathscr{S}^{\cdot}$ on $\mathbb{M}$ by

$$\begin{array}{l} \mathbb{M}; \\ \end{array} \begin{array}{ccccccccccccc} \cdots & \longrightarrow & \mathscr{S}^{-1} & \longrightarrow & \mathscr{S}^{0} & \longrightarrow & \mathscr{S}^{1} & \longrightarrow & \mathscr{S}^{2} & \longrightarrow & \mathscr{S}^{3} & \longrightarrow & \cdots \\ & & \| & & \| & & \| & & \| & & & & \\ \cdots & \longrightarrow & 0 & \longrightarrow & \mathcal{O} & \xrightarrow{\nabla^{B}_{A}} & \mathcal{O}^{B'}_{A}[-1]' & \xrightarrow{\nabla^{AA'}} & \mathcal{O}^{(A'B')}[-1][-2]' & \longrightarrow & 0 & \longrightarrow & \cdots \end{array}$$

then the spectral sequence of Theorem 4.1 is exactly the second hypercohomology spectral sequence of $\mathscr{S}^{\cdot}$ (cf. Swan [54], or the spectral sequence of a differential sheaf in Bredon [5] or Godement [14]). Thus, the interpretation of $H^r(U', \mu^{-1}\mathcal{O})$ as data on U is, in general, given by the isomorphism

$$H^r(U', \mu^{-1}\mathcal{O}) \cong \mathbb{H}^r(U, \mathscr{S}^{\cdot}). \tag{7.11}$$

Now define a second complex $\mathscr{R}^{\cdot}$ by

$$\begin{array}{l} \mathbb{M}; \\ \end{array} \begin{array}{ccccccccccccc} \cdots & \longrightarrow & \mathscr{R}^{1} & \longrightarrow & \mathscr{R}^{2} & \longrightarrow & \mathscr{R}^{3} & \longrightarrow & \mathscr{R}^{4} & \longrightarrow & \mathscr{R}^{5} & \longrightarrow & \cdots \\ & & \| & & \| & & \| & & \| & & & & \\ \cdots & \to & 0 & \longrightarrow & \mathcal{O}_{(AB)}[-1]' & \xrightarrow{\nabla^{B}_{B}} & \mathcal{O}^{B'}_{A}[-1][-2]' & \xrightarrow{\nabla^{A}_{B}} & \mathcal{O}[-2][-2]' & \to & 0 & \longrightarrow & \cdots \end{array}$$

Identifying the various spinor sheaves in terms of holomorphic forms, we can complete (7.6) to the exact sequence of complexes

$$\mathbb{M};\ 0 \longrightarrow \mathscr{R}^{\cdot} \longrightarrow \Omega^{\cdot} \longrightarrow \mathscr{S}^{\cdot} \longrightarrow 0$$

and hence obtain a long exact sequence on hypercohomology

$$\cdots \longrightarrow \mathbb{H}^1(U, \Omega^{\cdot}) \longrightarrow \mathbb{H}^1(U, \mathscr{S}^{\cdot}) \longrightarrow \mathbb{H}^2(U, \mathscr{R}^{\cdot}) \longrightarrow \mathbb{H}^2(U, \Omega^{\cdot}) \longrightarrow \cdots \tag{7.12}$$

Since $\Omega^{\cdot}$ is a resolution of $\mathbb{C}$,

$$\mathbb{H}^r(U, \Omega^{\cdot}) = H^r(U, \mathbb{C}). \tag{7.13}$$

The hypercohomology group $\mathbb{H}^2(U, \mathscr{R}^{\cdot})$ is easily calculated by the first hypercohomology spectral sequence, or from the definition of hypercohomology, and we find that

$$\mathbb{H}^2(U, \mathscr{R}^{\cdot}) = \Gamma(U, \mathscr{Z}'_{-2}), \quad \mathbb{H}^1(U, \mathscr{R}^{\cdot}) = 0 \tag{7.14}$$

Substituting (7.11), (7.13) and (7.14) into (7.12) we obtain the exact sequence of (cf. Eastwood [9])

$$0 \longrightarrow H^1(U, \mathbb{C}) \longrightarrow H^1(U', \mu^{-1}\mathcal{O}) \longrightarrow \Gamma(U, \mathscr{Z}'_{-2}) \longrightarrow H^2(U, \mathbb{C}) \longrightarrow \cdots \tag{7.15}$$

and $\mathscr{P}$ may be defined as the composition

$$H^1(U'', \mathcal{O}) \xrightarrow{\mu^*} H^1(U', \mu^{-1}\mathcal{O}) \longrightarrow \Gamma(U, \mathscr{Z}'_{-2}).$$

It is now easy to see when

$$\mathscr{P}: H^1(U'', \mathcal{O}) \longrightarrow \Gamma(U, \mathscr{Z}'_{-2})$$

is an isomorphism. Firstly, there are the usual topological conditions on $\mu: U' \to U''$ ensuring that μ^* be an isomorphism (Theorem 3.2). Secondly, (7.15) shows that $H^1(U', \mu^{-1}\mathcal{O}) \to \Gamma(U, \mathscr{Z}'_{-2})$ is an isomorphism if $H^1(U, \mathbb{C})$ and $H^2(U, \mathbb{C})$ vanish. Thus, the topological obstructions to existence and uniqueness of potentials remain whereas the analytic condition has disappeared. This suggests that perhaps

$\mathbb{H}^1(U, \mathscr{S}^\cdot)$ should be taken as a definition of potential modulo gauge freedom in the case of holomorphic forms.

8. A Direct Approach to Left-handed Fields

In Wells [53] a cohomological method was described for producing a left-handed field of helicity $-n/2$ directly from an element of $H^1(U'', \mathcal{O}(n-2))$. For this range of homogeneities the integral formulae (cf. Penrose [34]) involve differentiation in the integrand and this shows up in the cohomological method by using powers of the conormal bundle of a line in U'' to try to expand an element of $H^1(U'', \mathcal{O}(n-2))$ as a power series about the line. It is therefore quite straightforward to see that when an integral representation is possible, it agrees with this more abstract approach. What is not so clear is that this approach agrees with that described in the previous section of the present paper. In this section we shall first review the construction of Wells [53], describing it invariantly[10] with the aid of the spinor sheaves of Sect. 2. Then we shall use the process of *raising and lowering helicity* (cf. Eastwood [10] and Penrose [33], [39]) to show that it does indeed agree with the twistor description via potentials given in Sect. 7. This also follows by comparison of the constructions given in Penrose [37] and Ward [49].

Consider the bundle $\mathcal{O}^A(1)(:=\mathcal{O}(1)\otimes\mathcal{O}^A)$ on $\mathbb{P}\times\mathbb{M}$. The fibre of this bundle over $(L_1, L_2)\in\mathbb{P}\times\mathbb{M}$ is $L_1^*\otimes(\mathbb{T}/L_2)$. Choose $\ell\in L_1$. Let ℓ^* denote the dual element of L_1^* (i.e. $\ell^*(\ell)=1$), and let $[\ell]$ denote the image of ℓ under the quotient $\mathbb{T}\to\mathbb{T}/L_2$. Then $\ell^*\otimes[\ell]\in L_1^*\otimes(\mathbb{T}/L_2)$ is independent of the choice of ℓ and hence defines a section of $\mathcal{O}^A(1)$ which we will call ω^A. Observe that $\omega^A(L_1, L_2)=0$ if and only if $L_1\subset L_2$ and indeed ω^A defines $\mathbb{F}$ without multiplicitly as a subvariety of $\mathbb{P}\times\mathbb{M}$. In other words, if $\mathscr{J}$ denotes the ideal sheaf of $\mathbb{F}$ as a subvariety of $\mathbb{P}\times\mathbb{M}$, then

$$\begin{array}{rccc} & \mathcal{O}_A(-1) & \longrightarrow & \mathscr{J} \\ \mathbb{P}\times\mathbb{M}; & \cup & & \cup \\ & f_A & \longmapsto & f_A\omega^A \end{array}$$

is surjective. This also gives the following exact sequence

$$\begin{array}{rccccccc} \mathbb{P}\times\mathbb{M}; & 0 & \longrightarrow & \mathscr{J}^2 & \longrightarrow & \mathscr{J} & \longrightarrow & {}_{\mathbb{F}}\mathcal{O}_A(-1) \longrightarrow 0 \\ & & & & & \cup & & \cup \\ & & & & & f_A\omega^A & \longmapsto & f_A|_{\mathbb{F}} \end{array} \tag{8.1}$$

which identifies the conormal bundle $\mathscr{J}/\mathscr{J}^2$ as $\mathcal{O}_A(-1)$ on $\mathbb{F}$. Here, we have allowed ourselves to write ${}_X\mathcal{O}$ instead of $\mathcal{O}_X$ to avoid confusion with spinor indices. Now suppose U is an open Stein subset of $\mathbb{M}$ and consider the above construction

10 The construction itself is manifestly invariant but we repeat it here to show how it fits in with the spinor sheaves of Sect. 2. The use of the ideal sheaves in what follows is a device to make sense of $\partial/\partial\omega^A$ in an invariant manner. Knowing $\partial/\partial\omega^A$ is the same as knowing the location of I, the line in $\mathbb{P}$ corresponding to the vertex of the null-cone at infinity of $\mathbb{M}$. If this is fixed then it is easy to check that the method in Wells [53] agrees with that described by Penrose [37].

applied to U' as a subvariety of $U'' \times U$. We claim that

$$H^1(U'' \times U, \mathscr{J}^n(n-2)) \longrightarrow H^1(U'' \times U, \mathscr{O}(n-2)) \tag{8.2}$$

is an isomorphism. The general case is no more difficult than the helicity $-1\,(n=2)$ case so, for notational simplicity, we will restrict our proofs to this case. From the exact sequence of sheaves

$$U'' \times U;\ 0 \longrightarrow \mathscr{J} \longrightarrow \mathscr{O} \longrightarrow {}_{U'}\mathscr{O} \longrightarrow 0$$

we obtain an exact sequence on cohomology

$$\Gamma(U'' \times U, \mathscr{O}) \longrightarrow \Gamma(U', \mathscr{O}) \longrightarrow H^1(U'' \times U, \mathscr{J}) \longrightarrow H^1(U'' \times U, \mathscr{O}) \longrightarrow H^1(U', \mathscr{O}).$$

The holomorphic functions on $U'' \times U$ and U' are the holomorphic functions on U so $\Gamma(U'' \times U, \mathscr{O}) \to \Gamma(U', \mathscr{O})$ is surjective. Also, by Proposition 2.2, $H^1(U', \mathscr{O}) \cong H^1(U, \mathscr{O}) = 0$, since U is Stein. Hence

$$H^1(U'' \times U, \mathscr{J}) \longrightarrow H^1(U'' \times U, \mathscr{O})$$

is an isomorphism. Now consider the exact sequence (8.1) on $U'' \times U$. The corresponding long exact sequence on cohomology includes the portion

$$\Gamma(U', \mathscr{O}_A(-1)) \longrightarrow H^1(U'' \times U, \mathscr{J}^2) \longrightarrow H^1(U'' \times U, \mathscr{J}) \longrightarrow H^1(U', \mathscr{O}_A(-1))$$

and Proposition 2.2 shows that the first and last terms vanish. Hence we may conclude

$$H^1(U'' \times U, \mathscr{J}^2) \xrightarrow{\cong} H^1(U'' \times U, \mathscr{J}) \xrightarrow{\cong} H^1(U'' \times U, \mathscr{O})$$

as required. Intuitively, (8.2) being an isomorphism says that the information contained within an element of $H^1(U'' \times U, \mathscr{O}(n-2))$ depends only on normal derivatives to U' of order at least n. To evaluate the field we in effect just take the n^{th} normal derivative by factoring out $\mathscr{J}^{n+1}(n-2)$. This also has the required effect of restricting to U' since $\mathscr{J}^n(n-2)/\mathscr{J}^{n+1}(n-2)$ is supported on U'. It follows from the exact sequence (cf. 8.1)

$$\begin{array}{ccccccccc}
\mathbb{P} \times \mathbb{M};\ 0 & \longrightarrow & \mathscr{J}^{n+1} & \longrightarrow & \mathscr{J}^n & \longrightarrow & {}_{U'}\mathscr{O}_{(AB\cdots D)}(-n) & \longrightarrow & 0 \\
 & & & & \cup & & \cup & & \\
 & & & & f_{AB\cdots D}\omega^A\omega^B \ldots \omega^D & \longmapsto & f_{AB\cdots D}|_{U'} & &
\end{array}$$

that $\mathscr{J}^n(n-2)/\mathscr{J}^{n+1}(n-2)$ may be canonically identified with $\mathscr{O}_{(AB\cdots D)}(-2)$ on U'. The $\mathscr{P}$-transform in Wells [53] is defined as the composition

$$\begin{array}{ccccc}
H^1(U'', \mathscr{O}(n-2)) & \longrightarrow & H^1(U'' \times U, \mathscr{O}(n-2)) & \xleftarrow{\cong} & H^1(U'' \times U, \mathscr{J}^n(n-2)) \\
 & \searrow & & & \downarrow \\
 & & & & H^1(U', \mathscr{O}_{(AB\cdots D)}(-2)) \\
 & & & & \| \\
 & & & & \Gamma(U, \mathscr{O}_{(AB\cdots D)}[-1]')
\end{array} \tag{8.3}$$

where the last isomorphism follows from Proposition 2.2. To show that (8.2) was an isomorphism we assumed that U was Stein (though we only used $H^1(U, \mathscr{O}) = 0$ in case $n = 2$). This indicates strongly that the construction is really via potentials especially as, in case $n = 1$, the assumption is unnecessary. To get around this

problem we simply perform the transform (8.3) locally. It is clear that these local transforms will agree on overlaps and hence patch together to give a well-defined transform for any U. An alternative approach is to apply the argument to a fixed line L_z, for $z \in U$, as a subvariety of U'' instead of U' as a subvariety of $U'' \times U$. This is the approach that we shall adopt to show that $\mathscr{P}$ defined by (8.3) agrees (apart from a combinatorial constant factor) with $\mathscr{P}$ defined by Theorem 7.1.

Again we will restrict our discussion to the helicity -1 case. Choose a point of U. We shall call this point the origin and introduce coordinates $z^{AA'}$ as in Sect. 1 so that this origin really is the origin of $\mathbb{M}^I$. Also introduce coordinates $(\omega^A, \pi_{A'})$ on $\mathbb{T}$ as in Sect. 1. Then the line L in $\mathbb{P}$ corresponding to the origin is defined by the vanishing of ω^A. If $\mathscr{J}$ denotes the ideal sheaf of L as a subvariety of U'' then the value of $\mathscr{P}$ at the origin according to (8.3) is given by the composition

$$H^1(U'', \mathscr{O}) \xleftarrow{\cong} H^1(U'', \mathscr{J}^2) \longrightarrow H^1(L, \mathscr{J}^2/\mathscr{J}^3) \cong H^1(L, {}_L\mathscr{O}_{(AB)}(-2)) \quad (8.4)$$

where the last space is regarded as the fibre of $\mathscr{O}_{(AB)}[-1]'$ over the origin. To compare this with Sect. 7 we proceed as follows. Since $\mathscr{J}^2$ is generated by $\omega^A\omega^B$ we have the exact sequence

$$0 \longrightarrow \mathscr{K} \longrightarrow \mathscr{O}_{(AB)}(-2) \xrightarrow{\omega^A\omega^B} \mathscr{J}^2 \longrightarrow 0 \quad (8.5)$$

where $\mathscr{K}$ is some coherent analytic sheaf. Consider the following commutative diagram

$$\begin{array}{ccccc} H^1(U'', \mathscr{O}_{(AB)}(-2)) & \xrightarrow{\omega^A\omega^B} & H^1(U'', \mathscr{J}^2) & \xrightarrow{\cong} & H^1(U'', \mathscr{O}) \\ \downarrow & & \downarrow & & \\ H^1(L, {}_L\mathscr{O}_{(AB)}(-2)) & \cong & H^1(U'', \mathscr{J}^2/\mathscr{J}^3) & & \end{array} \quad (8.6)$$

Tracing this diagram from $H^1(U'', \mathscr{O})$ to $H^1(L, {}_L\mathscr{O}_{(AB)}(-2))$ is the proposed construction (8.4) whereas the restriction $H^1(U'', \mathscr{O}_{(AB)}(-2)) \to H^1(L, {}_L\mathscr{O}_{(AB)}(-2))$ is the construction of the value of a field at the origin according to Sect. 6. The interpretation of

$$\omega^A\omega^B : H^1(U'', \mathscr{O}_{(AB)}(-2)) \longrightarrow H^1(U'', \mathscr{O})$$

on the corresponding massless fields (via Sect. 6 and Sect. 7) can be explicitly computed (Eastwood [10] and Penrose [39]) and turns out to be

$$\phi_{AB} \longmapsto -[z^{CA'}z^{DB'}\nabla_{AA'}\nabla_{BB'}\phi_{CD} + 4z^{CA'}\nabla_{A'(A}\phi_{B)C} + 2\phi_{AB}].$$

Hence, the value of the field at the origin is simply multiplied by -2. If $\omega^A\omega^B$: $H^1(U'', \mathscr{O}_{(AB)}(-2)) \to H^1(U'', \mathscr{J}^2)$ were surjective, then we would be able to conclude from (8.6) that, apart from the factor of -2, the method of Sect. 7 agreed with (8.4). Then, all the conclusions of Sect. 7 would carry over. If $H^2(U'', \mathscr{K}) = 0$ then (8.5) would imply the required surjectivity. For suitable U this is true (cf. Andreotti and Grauert [1]) but we can circumvent this more complicated argument as follows. The value of the field at the origin is unchanged if we shrink U (and hence U''). We claim it is always possible to shrink U to V, say, so that the corresponding map $H^2(U'', \mathscr{K}) \to H^2(V'', \mathscr{K})$ is the zero map. This would certainly suffice to complete the comparison and is a consequence of the following lemma.

Lemma 8.1. *Any line L in $\mathbb{P}$ has arbitrarily small open neighbourhoods W with $H^2(W, \mathscr{S}) = 0$ for all coherent analytic sheaves $\mathscr{S}$.*

Proof. Choose two standard coordinate patches on $\mathbb{P}$ whose union contains L. In each of these patches there are arbitrarily small Stein sets containing the part of L lying in the patch. Thus a neighbourhood W may be chosen which is a union of two Stein sets. The Mayer–Vietoris theorem and Cartan's theorem B (cf. Hörmander [20] or Gunning and Rossi [18]) show that $H^2(W, \mathscr{S}) = 0$ for all coherent analytic sheaves $\mathscr{S}$. □

9. Background Coupled Fields

In [46], [47] Ward gave a method involving the "twisted photon" for describing certain background coupled fields by means of twistor integral formulae. In this section we shall show how the general machine of Sect. 4 may be combined with the results of Sect. 7 to interpret this result cohomologically.

Suppose $\Phi_A^{A'}$ is a potential for a left-handed photon (Maxwell field) on U in the sense of Sect. 7. Define a differential operator $D_A^{A'}$ on U by

$$D_A^{A'} := \nabla_A^{A'} + 2\pi i \Phi_A^{A'} : \mathcal{O}_{(A'B'\cdots D')}[-1] \longrightarrow \mathcal{O}_{A(B'\cdots D')}[-1][-1]'$$

Then (cf. Gasiorowicz [13]), a holomorphic right-handed massless field on U *minimally coupled into the electromagnetic background defined by* $\Phi_A^{A'}$ is a section of $\mathcal{O}_{(A'B'\cdots D')}[-1]$ satisfying

$$D_A^{A'} \phi_{A'B'\cdots D'} = 0$$

If there are n indices, then the field is said to be of helicity $n/2$. We will denote the sheaf of such by $\mathscr{Z}_n(\Phi)$. If $\Phi_A^{A'} = 0$ then, of course, we just have $\mathscr{Z}_n$ as before in (5.1). Also, there is the background coupled version of (5.1)′ which will be denoted by $\mathscr{Z}_n'(\Phi)$.

If we replace $\Phi_A^{A'}$ by an equivalent potential (cf. (7.3))

$$\Phi_A^{A'} + \nabla_A^{A'}\gamma,$$

then the background coupled fields are essentially undisturbed since

$$\nabla_A^{A'} + 2\pi i \Phi_A^{A'} = e^{2\pi i\gamma}(\nabla_A^{A'} + 2\pi i(\Phi_A^{A'} + \nabla_A^{A'}\gamma))e^{-2\pi i\gamma}$$

Hence, the potential is only important up to the usual gauge freedom.

In Sect. 7 we described how left-handed photons may be represented under suitable conditions by elements of $H^1(U'', \mathcal{O})$. Since this representation was via potentials modulo gauge freedom we may expect it to be useful in describing electromagnetically coupled fields. This is indeed the case. The representation of a left-handed photon as an element of $H^1(U'', \mathcal{O})$ is called the *passive description*. The *active description* is in terms of a line-bundle over U'', namely the *twisted photon* (Ward [46], [47]).[11] To obtain this bundle from the passive description

11 The "active" and "passive" terminology used here differs from that used in Penrose and MacCallum [32]

consider the short exact sequence of sheaves

$$\mathbb{P};\ 0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O} \xrightarrow{\exp 2\pi i} \mathcal{O}^* \longrightarrow 0$$

where $\mathbb{Z}$ denotes the constant sheaf of integers and $\mathcal{O}^*$ the sheaf of nowhere-vanishing holomorphic functions. The long exact sequence on cohomology contains the portion

$$H^1(U'', \mathbb{Z}) \longrightarrow H^1(U'', \mathcal{O}) \longrightarrow H^1(U'', \mathcal{O}^*) \xrightarrow{c} H^2(U'', \mathbb{Z}) \qquad (9.1)$$

and the *twisted photon* L is defined as the image of the passive description under the exponential map

$$H^1(U'', \mathcal{O}) \longrightarrow H^1(U'', \mathcal{O}^*).$$

The possible line-bundles obtainable by this process are, by (9.1), precisely those which are topologically trivial (i.e. $c(L) = 0$). Moreover, recall (Sect. 7) that there are conditions on U which must be satisfied for the description of a left-handed Maxwell field as an element of $H^1(U'', \mathcal{O})$ to be valid including

a) $$H^1(U, \mathbb{C}) = 0, \qquad (9.2)$$

b) $\mu : U' \to U''$ is elementary.

It is straightforward to show that conditions (9.2) imply $H^1(U'', \mathbb{Z}) = 0$ so that left-handed photons are precisely in $1-1$ correspondence with topologically trivial line-bundles on U''. This is connected with "charge integrality" (see Penrose [36]).

We need not impose such restrictions, however, just to obtain a cohomological description of background coupled fields. We need only suppose that we are given a twisted photon (i.e. a topologically trivial line-bundle on U'') which happens to correspond to a potential on U in a sense which may be made precise as follows. Recall the exact sequence (1.3)

$$\mathbb{F};\ 0 \longrightarrow \mu^{-1}\mathcal{O} \longrightarrow \mathcal{O} \xrightarrow{d_*} \Omega^1_\mu \xrightarrow{d_*} \Omega^2_\mu \longrightarrow 0$$

Define a sheaf $\mathcal{K}$ on $\mathbb{F}$ by

$$\mathcal{K} := \ker d_\mu : \Omega^1_\mu \longrightarrow \Omega^2_\mu$$

Then we obtain an exact sequence

$$\begin{array}{ccccc} 0 \longrightarrow \Gamma(U', \mathcal{K}) \longrightarrow \Gamma(U', \Omega^1_\mu) & \xrightarrow{d_*} & \Gamma(U', \Omega^2_\mu) \\ \| & & \| \\ \Gamma(U, \mathcal{O}^{A'}_{A}[-1]') & \xrightarrow{\nabla^{AB}} & \Gamma(U, \mathcal{O}^{(A'B')}[-1][-2]') \end{array}$$

so $\Gamma(U', \mathcal{K})$ is isomorphic to the space of potentials on U. The exact sequence

$$\mathbb{F};\ 0 \longrightarrow \mu^{-1}\mathcal{O} \longrightarrow \mathcal{O} \xrightarrow{d_*} \mathcal{K} \longrightarrow 0$$

gives rise to an exact sequence on cohomology

$$\begin{array}{ccccccccc} \cdots \longrightarrow \Gamma(U', \mathcal{O}) & \xrightarrow{d_*} & \Gamma(U', \mathcal{K}) & \xrightarrow{\delta} & H^1(U', \mu^{-1}\mathcal{O}) & \longrightarrow & H^1(U', \mathcal{O}) & \longrightarrow & \cdots \\ \| & & \| & & & & & & \\ \Gamma(U, \mathcal{O}) & \xrightarrow{\nabla^A_A} & \{\text{potentials on } U\}. & & & & & & \end{array} \qquad (9.3)$$

Thus, an element in $H^1(U', \mu^{-1}\mathcal{O})$ gives rise to a potential on U if and only if its image in $H^1(U', \mathcal{O})$ is zero. Furthermore, the potential is unique up to the usual gauge freedom. We shall say that a *potential* $\Phi_{\mathcal{A}}^{\mathcal{A}'}$ *corresponds to a twisted photon* L if and only if we can find $\omega \in H^1(U'', \mathcal{O})$ so that $L = \exp 2\pi i\omega$ and $\mu^*\omega = \delta\Phi$, where $I(\Phi) = \Phi_{\mathcal{A}}^{\mathcal{A}'}$ under the direct image isomorphism I in (9.3). If the hypotheses of Theorem 7.1 are satisfied then this agrees with taking the passive description of that theorem and then exponentiating.

The cohomological description of coupled fields is obtained by replacing H^n by $L \otimes H^n$ in the usual description of uncoupled fields (cf. Theorem 5.1).

Theorem 9.1. *Suppose that L is a twisted photon corresponding to a potential $\Phi_{\mathcal{A}}^{\mathcal{A}'}$ on U for a left-handed Maxwell field. Then, for $n \geqq 1$, there is a canonical linear transformation*

$$\mathcal{P} : H^1(U'', \mathcal{O}(L)(-n-2)) \longrightarrow \Gamma(U, \mathcal{X}'_n(\Phi)).$$

If $\mu : U' \to U''$ has connected fibres, then $\mathcal{P}$ is injective and if $\mu : U' \to U''$ is elementary, then $\mathcal{P}$ is an isomorphism.

Proof. The proof follows that of Theorem 5.1 and $\mathcal{P}$ is constructed by first applying

$$\mu^* : H^1(U'', \mathcal{O}(L)(-n-2)) \longrightarrow H^1(U', \mu^{-1}\mathcal{O}(L)(-n-2)).$$

It now suffices to show that

$$H^1(U', \mu^{-1}\mathcal{O}(L)(-n-2)) \cong \Gamma(U, \mathcal{X}'_n(\Phi)).$$

We do this by using Theorem 4.1, as usual. Hence, we must compute $H^q(U', \Omega^p_\mu(L)(-n-2))$ and the differentials

$$d_\mu : H^q(U', \Omega^p_\mu(L)(-n-2)) \longrightarrow H^q(U', \Omega^{p+1}_\mu(L)(-n-2)). \tag{9.4}$$

Since L corresponds to a potential $\Phi_{\mathcal{A}}^{\mathcal{A}'}$ we have the following commutative diagram (cf. (9.3)).

$$\begin{array}{ccccccc}
\Gamma(U', \mathcal{K}) & \longrightarrow & H^1(U', \mu^{-1}\mathcal{O}) & \longrightarrow & H^1(U', \mathcal{O}) & \longrightarrow & H^1(U', \mathcal{O}^*) \\
\cup & & \cup & & \cup & & \cup \\
\Phi & \longmapsto & \mu^*\omega & \longmapsto & 0 & \longmapsto & 0 \\
 & & \updownarrow & & & & \updownarrow \\
 & & \omega & & \longmapsto & & L \\
 & & \cap & & & & \cap \\
 & & H^1(U'', \mathcal{O}) & & \longrightarrow & & H^1(U'', \mathcal{O}^*)
\end{array} \tag{9.5}$$

Thus, μ^*L is trivial as a holomorphic line bundle on U', whence

$$\Omega^p_\mu(L)(-n-2) \cong \Omega^p_\mu(-n-2).$$

Hence, the terms in the E_1-level of the spectral sequence of Theorem 4.1 are exactly as in the case of Theorem 5.1 where L is trivial. The only difference is in the differentials where we expect $\nabla_{\mathcal{A}}^{\mathcal{A}'}$ to be replaced by $D_{\mathcal{A}}^{\mathcal{A}'}$ throughout. The differentials are given by (9.4) and so it suffices to show that if we trivialize μ^*L in a natural manner (using $\Phi_{\mathcal{A}}^{\mathcal{A}'}$), then we have the following commutative diagram

 Cohomology and massless fields. *Commun. Math. Phys.* **78** (1980/81), no. 3, 305–51. With M. G. Eastwood and R. O. Wells, Jr.

$$\begin{array}{ccc} \Omega^0_\mu(L) & \xrightarrow{d_\mu} & \Omega^1_\mu(L) \\ \| & & \| \\ \mathbb{F}: \quad \mathcal{O} & \longrightarrow & \mathcal{O}_{A'}(1)[-1]' \\ \cup & & \cup \\ f & \longmapsto & \pi_{A'} D^{A'}_A f \end{array} \tag{9.6}$$

for this is the appropriate generalization of Lemma 1.2. The trivialization of μ^*L is given implicitly by (9.5). To make this explicit we use the Čech definition of cohomology. Let $\{U_j\}_{j\in J}$ be an open cover of U'' such that ω has a representation as a Čech cocycle:

$$\omega_{jk} \in \Gamma(U_j \cap U_k, \mathcal{O}).$$

$\mu^*\omega$ is then defined by the cocycle

$$\omega_{jk} \circ \mu \in \Gamma(\mu^{-1}(U_j) \cap \mu^{-1}(U_k), \mu^{-1}\mathcal{O}),$$

with respect to the cover $\{\mu^{-1}(U_j)\}$. From (9.5), $\delta\Phi = \mu^*\omega$, i.e. there is:

1) An open cover $\{V_j\}_{j\in J}$ of U' (after perhaps redefining the indexing of $\{U_j\}$) so that $V_j \subseteq \mu^{-1}(U_j)$)
2) $h_j \in \Gamma(V_j, \mathcal{O})$ so that $\begin{cases} d_\mu h_j = \Phi|_{V_j} \\ \omega_{jk} \circ \mu = h_j - h_k \quad \text{on } V_j \cap V_k. \end{cases}$

Hence, from this last equation

$$(\exp 2\pi i \omega_{jk} \circ \mu) \exp 2\pi i h_k = \exp 2\pi i h_j \text{ on } V_j \cap V_k$$

But $\{\exp 2\pi i \omega_{jk} \circ \mu\}$ are transition functions for the bundle μ^*L so $\{\exp 2\pi i h_j\}$ is a nowhere-vanishing section of μ^*L which therefore trivializes it:

$$\begin{array}{ccc} \mathcal{O} = \mathcal{O}(\mu^*L) := \Omega^0_\mu(L) \\ \mathbb{F}: \quad \cup \qquad\qquad \cup \\ f \quad \longmapsto \{f \exp 2\pi i h_j\}. \end{array}$$

Now we apply $d_\mu : \Omega^0_\mu(L) \to \Omega^1_\mu(L)$ to this section:

$$\begin{aligned} d_\mu(f \exp 2\pi i h_j) &= d_\mu f \exp 2\pi i h_j + f(2\pi i d_\mu h_j) \exp 2\pi i h_j \\ &= [(d_\mu + 2\pi i \Phi) f] \exp 2\pi i h_j. \end{aligned}$$

Identifying $\Omega^1_\mu(L)$ with $\mathcal{O}_{A'}(1)[-1]'$ using the same trivialization, we obtain (9.6) as required. □

The above theorem interprets $H^1(U', \mathcal{O}(L)(k))$ for $k < -2$. For $k = -2$, an analogue of Theorem 6.1 holds, and for $k > -2$, an analogue of Theorem 7.1. All that is necessary is to replace $\nabla^{A'}_A$ by $D^{A'}_A$ in the statements and proofs of the theorems. Note that $\nabla^{A(A'} \Phi^{B')}_A = 0$ is equivalent to saying that $D^{(B'}_A D^{A')A} = 0$.

Hence, replacing the operator $\nabla^{A'}_A$ by the deformed operator $D^{A'}_A$ is equivalent to deforming the trivial bundle to obtain L. This point of view, of regarding L as obtained by deforming the trivial bundle, is how the twisted photon was originally constructed by Ward [46], [47] (though the definition given there has a twist of -1 also). Actually, it is best to regard $D^{A'}_A$ as a connection for the trivial

line bundle on U. The twisted photon construction then generalizes as a $1-1$ correspondence (provided $\mu: U' \to U''$ is elementary) between holomorphic line-bundles on U with anti-self-dual connection and holomorphic line-bundles on U'' which are trivial when restricted to lines L_z for $z \in U$. Theorem 9.1 is easily generalized to cover this. Indeed, the general case is immediate by patching the statement for trivial bundles.

For suitable U (e.g. convex in $\mathbb{M}^I$) we have $H^2(U'', \mathbb{Z}) \cong \mathbb{Z}$ so all line-bundles are of the form $L(k)$ for a suitable twisted photon L, and $c(L(k)) = k$ where c is the Chern class mapping: $H^1(U'', \mathcal{O}^*) \to H^2(U'', \mathbb{Z})$. Thus, the Chern class of a line-bundle V is simply related to the helicity of the field represented by $H^1(U'', \mathcal{O}(V))$. Specifically,

$$\text{helicity} = (-c(V) - 2)/2.$$

In [43] and [44] Ward also described how the twisted photon construction may be generalized to non-Abelian gauge groups. Left-handed Yang–Mills fields are then described by holomorphic vector-bundles on U''. If we take V to be one of these bundles in $H^1(U', \mathcal{O}(V))$, then we generate fields coupled to Yang–Mills potentials (see Rawnsley [42] and Hitchin [19]).

10. Further Remarks

The results of Sect. 5 through Sect. 9 are concerned with the interpretation of the first cohomology of U'' with coefficients in various analytic sheaves in terms of corresponding differential equations on U. The general $\mathscr{P}$-transform (Theorem 4.1), however, is not limited to first cohomology and we should ask what it says about $H^p(U'', \mathcal{O}(n))$ for p other than 1. We consider first $H^0(U'', \mathcal{O}(n)) = \Gamma(U'', \mathcal{O}(n))$, and we find, for any open set $U \subset \mathbb{M}$:

$$\mathscr{P}: \Gamma(U'', \mathcal{O}(n)) \xrightarrow{\cong} \ker \nabla_A^{E'} : \Gamma(U, \mathcal{O}^{(A'B'\cdots D')}) \longrightarrow \Gamma(U, \mathcal{O}_A^{(A'B'\cdots E')}), \quad \text{if } n \geqq 0.$$

$$\Gamma(U'', \mathcal{O}(n)) = 0, \quad \text{if } n \leqq -1.$$

For U sufficiently nice, (e.g. convex $U \subset \mathbb{M}^I$) one finds that U'' satisfies the hypotheses of Andreotti–Grauert's vanishing theorem: $H^q(U'', \mathscr{S}) = 0, q \geqq 2$, for any coherent sheaf $\mathscr{S}$ on U'', so the $\mathscr{P}$-transform of these higher degree cohomology groups would not yield too much information (cf. Lemma 8.1). One could readily compute what the image was, and it would in general be of the form [potentials]/[gauge freedom], and the $\mathscr{P}$-transform being zero would simply say that each potential was gauge equivalent to zero in this case. We remark that $H^3(U'', \mathscr{S}) = 0$ for *any* U'' and a coherent sheaf $\mathscr{S}$ on U'' (Siu [44]; Malgrange [25], in the case of a locally free sheaf as we have), so there are no nontrivial fields generated by these cohomology groups for any geometry.

It is easy to show that every section of $\mathcal{O}(n)$ over U'' extends uniquely to a section of $\mathcal{O}(n)$ over all of $\mathbb{P}$. Thus, we have canonical isomorphisms

$$\Gamma(U'', \mathcal{O}(n)) = \Gamma(\mathbb{P}, \mathcal{O}(n)) = \begin{cases} \odot^n \mathbb{T}^* & \text{if } n \geqq 0. \\ \odot & \text{if } n \leqq -1. \end{cases}$$

In particular ($n = 1$), this shows that a dual twistor $(\eta_A, \xi^{A'})$ may be regarded as a solution of the equation

$$\nabla_A^{(B'}\zeta^{A')} = 0 \tag{10.1}$$

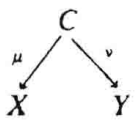
Indeed, it is easy to show that the general solution of (10.1) is

$$\xi^{A'} = \xi^{A'} + iz^{AA'}\eta_A$$

for constant spinors $\xi^{A'}$ and η_A. Thus, (10.1) may be regarded as defining $\mathbb{I}^*$. This equation is called the *dual twistor equation*. Similarly, twistors may be regarded as solutions of the *twistor equation* (cf. Penrose [28])

$$\nabla_{A'}^{(B}\theta^{A)} = 0. \tag{10.2}$$

We conclude with some general remarks concerning the methods of this paper. It is clear that our approach is not limited only to the usual twistor correspondence (1.1). All we need is some general correspondence



(10.3)

with ν having nice fibres. In case of (1.1) the fibres of ν are Riemann spheres so that any vector-bundle splits as a direct sum of line-bundles which are characterized by their Chern classes (cf. Grothendieck [17]). Thus it is possible to identify the appropriate direct images and produce a transform starting with any vector bundle V over a portion of $\mathbb{P}$. In general, things may be more complicated but we can still handle many interesting special cases of (10.3) where the fibres of ν are higher-dimensional projective spaces (or flag manifolds). For example, it is possible to give a cohomological interpretation of the integral formulae described by Hughston [22], in connection with *generalized twistors*. Another possibility is to replace U'', U', U used extensively in this paper by $X \subset \mathbb{P}$, $Y \subset \mathbb{M}$, and $C \subset \mathbb{F}$ so that (10.3) is a surjective double fibration. It need not be the case that $X = Y''$, for instance, and more general phenomena can occur.[12] A particularly interesting generalization is to the case when (1.1), or rather (1.2), is *deformed*. The resulting deformation of twistor space is called the *non-linear graviton construction* (Penrose [35]) and the corresponding space-time is automatically right-flat. The methods of this paper give a cohomological interpretation of the integral formulae of Penrose and Ward [41]. In this right-flat space-time the twistor description of left-handed fields via potentials is no longer equivalent to the field equations. This suggests that the correct description of left-handed fields in a right-flat background is as potentials modulo gauge freedom (see Hitchin [19] and Penrose and Ward [41] for details). There are a least two more examples where the methods of this paper apply. One is to obtain a cohomological understanding of the twistor integral formulae for massive fields (cf. Penrose [33] and Hughston [23]). In

12 cf. Hodges A.P.: Cohomological wave functions. Twistor Newsletter 7, 18, Oxford University, June 1978

the case of two twistors see Eastwood [11]. The other example is in the application of twistor theory to the classification of instantons given by Atiyah, Hitchin, Drinfeld and Manin [2]. If we identify $\mathbb{T}$ with $\mathbb{H}^2$, the 2-dimensional quaternionic vector-space then we obtain an inclusion

$$S^4 = \mathbb{P}_1(\mathbb{H}) \subset \mathbb{M}.$$

In [2] the $\mathscr{P}$-transform is used to identify $H^1(U'', \mathcal{O}(E)(-2))$ with solutions of a certain differential equation (the Yang-Mills coupled wave-equation) on an open subset U of S^4, where E is an "instanton bundle" on $\mathbb{P}$. To establish this result using the methods of this paper we thicken S^4 a little inside $\mathbb{M}$, using the ellipticity of the wave-equation to extend the solution to this thickening. Theorem 3.2 then applies to show μ^* is an isomorphism and the argument continues as in Sect. 6.

Finally, we should remark that the seemingly abstract arguments of this paper can be made quite explicit to give integral formulae for $\mathbb{P}$ and its inverse. The original integral formulae for constructing massless fields (cf. Penrose [29] etc.) really use a Čech representation for the cohomology. This is very useful as it allows the residue theorem of one-variable complex analysis to be applied to produce specific examples of massless fields (especially the "elementary states" (cf. Penrose [33])). For more general arguments this method is not so good, since the choice of contour depends on the choice of covering and can therefore become quite involved (using "branched-contours"). To avoid this one can use the Dolbeault representation of cohomology and integrate over the whole lines L_z in $\mathbb{P}$ rather than contours upon these "lines" (really Riemann spheres). This method is described by Woodhouse [55] and Wells [52]. Both methods really come down to identifying the cohomology $H^1(\mathbb{P}_1(\mathbb{C}), \mathcal{O}(k))$ by means of Serre duality [43]. This integration has been replaced by taking the first direct image $\nu^1_* \mathcal{O}(k)$ in (2.11). To describe the inverse transform $\mathscr{P}^{-1}$ explicitly we must use the integral formula in Lemma 3.1 to reverse μ^* as in Theorem 3.2, together with an explicit method of inverting the direct image ν^1_*. Both of these steps involve some choice but this is only to be expected as the Dolbeault form produced in this way is supposed to be only a representative (one of many) for a cohomology class. A particularly natural way of inverting ν^1_* is to choose harmonic representatives for the cohomology of the fibres of ν. This is described explicitly by Woodhouse [55] for massless fields and by Rawnsley [42] for instanton bundles. For the potential description of left-handed fields (Sect. 7) integral formulae can be avoided for $\mathbb{P}$ since we are only using Liouville's theorem that $\mathbb{P}_1(\mathbb{C})$ supports only constant holomorphic functions (cf. Penrose and Ward [41]).

In our discussion of fields in $\mathbb{M}$ we have been careful to maintain a distinction between primed and unprimed weights. The physical significance of this distinction is not altogether clear, and its role played in this paper is not a crucial one. However, it may be pointed out that if a local rescaling of the metric is made, as was discussed in Sect. 2 (cf. (2.4), (2.5)), according to which $\varepsilon_{AB} \mapsto \Omega\varepsilon_{AB}$ and $\varepsilon_{A'B'} \mapsto \tilde{\Omega}\varepsilon_{A'B'}$ with Ω and $\tilde{\Omega}$ independent scalar functions on (subsets of) $\mathbb{M}$ then the field equations are *locally* invariant for helicity $\frac{1}{2}r$ (where r is a positive or negative integer) if the weights are taken to be $[-\frac{1}{4}r - \frac{1}{2}]\,[\frac{1}{4}r - \frac{1}{2}]'$. This works also for

$r = 0$ if the wave equation is taken in the form $(\Box + \frac{1}{6}R)\phi = 0$, where R is the scalar curvature, the metric being rescaled according to $g_{ab} \to \Omega\tilde{\Omega}g_{ab}$. The potential description of Sect. 7 is invariant for helicity $-\frac{1}{2}n (n > 1)$ if the weights are $[-\frac{1}{4}n]$ $[\frac{1}{4}n - 1]'$. When the ratio $\Omega/\tilde{\Omega}$ is constant over $\mathbb{M}$ (as is the case for the global $GL(\mathbb{T})$ transformations considered here) local conformal invariance depends only on the sum of the primed and unprimed weights (cf. (2.4)), and for massless fields the value -1 for this sum, as obtained in Sects. 5–8, agrees with the above local invariance.

The role of the more general conformal invariance that these fields exhibit when $\Omega/\tilde{\Omega}$ need not be constant is obscure, but the specific weights that arise, as mentioned above, seem to be related to the phenomenon of Grgin [15a] according to which it is only the fields of *odd* integral helicity that can satisfy the field equations globally on *real* compactified Minkowski space, assuming that they are to be represented as tensor fields in the ordinary way. Such fields of even integral helicity would require a twofold cover of this space and those of half-odd helicity, a fourfold cover. This arises from the fact that the fields change by a factor of $(-1)^{(1/4)r-1/2}$ across infinity. It is, however, not necessary to take these covering spaces if the fields are regarded as cross-sections of the appropriate twisted vector bundles (see Lerner [24a]). In fact, this interpretation is implicit in the descriptions of this paper. The definition of the spin-bundles $\mathcal{O}_{A'}$ and $\mathcal{O}_{A}$ that were introduced in Sect. 2 incorporate this twist in a natural way (cf. also [32]).

References

1. Andreotti, A., Grauert, H.: Théorèmes de finitude pour cohomologie des espaces complexes. Bull. Soc. Math. France **90**, 193–259 (1962)
2. Atiyah, M. F., Hitchin, N. J., Drinfeld, V. G., Manin, Yu. I.: Construction of instantons. Phys. Lett. **65A**, 185–187 (1978)
3. Atiyah, M. F., Ward, R.: Instantons and algebraic geometry. Commun. Math. Phys. **55**, 111–124 (1977)
4. Bateman, H.: The solution of partial differential equations by means of definite integrals. Proc. London Math. Soc. **(2)1**, 451–458 (1904)
5. Bredon, G. E.: Sheaf theory. New York, St. Louis, San Francisco, Toronto, London, Sydney: McGraw-Hill 1967
6. Dirac, P. A. M.: Relativistic wave equations. Proc. R. Soc. London Ser. A **155**, 447–459 (1936)
7. Drinfeld, V. G., Manin, Yu. I.: Instantons and sheaves on $\mathbb{CP}^3$. Functional Anal. Appl. **13**, 59–74 (1979), (Engl. trans. 124–134 (1979))
8. Eastwood, M. G.: Some cohomological arguments applicable to twistor theory. Twistor Newsletter **7**, 6–10 (1978)[13]
9. Eastwood, M. G.: Zero-rest-mass fields and topology. Twistor Newsletter **7**, 11–11 (1978)[13]
10. Eastwood, M. G.: On raising and lowering helicity. Twistor Newsletter **8**, 37–38 (1979)[13]
11. Eastwood, M. G.: Ambitwistors. Twistor Newsletter **9**, 55–58 (1979)
12. Fierz, M.: Über die relativistische Theorie kräftefreier Teilchen mit beliebigem Spin. Helv. Phys. Acta **12**, 3–37 (1939)
13. Gasiorowicz, S: Elementary particle physics. New York: Wiley 1966
14. Godement R.: Topologie algébrique et théorie des faisceaux. Paris: Hermann 1964

13 *Twistor Newsletter* is a series of informal reports distributed in preprint form from Oxford University. An edited version of issues 1–8 are published in: *Advances in Twistor theory* (eds. L.P. Hughston, R.S. Ward). Research Notes in Math 37. San Francisco, London, Melbourne: Pitman 1979

15. Gindikhin, S., Henkin, G.: Integral geometry for $\bar{\partial}$-cohomology in q-linearly concave domains in $\mathbb{CP}^n$ (Russian). Functional Anal. Appl. **12**, 6–23 (1978) (Engl. Trans. 247–261 (1979)
15a. Grgin, E.: A global technique for the study of spinor fields. Ph.D. Thesis, Department of Physics, Syracuse University 1966
16. Griffiths, P., Adams, J.: Topics in algebraic and analytic geometry. In: Math. Notes 13. Princeton: Princeton University Press 1974
17. Grothendieck, A.: Sur la classification des fibrés holomorphes sur la sphère de Riemann. Am. J. Math. **79**, 121–138 (1957)
18. Gunning, R. C., Rossi, H.: Analytic functions of several complex variables. Englewood Cliffs, N. J.: Prentice-Hall 1965
19. Hitchin, N. J.: Linear field equations on self-dual spaces. Proc. R. Soc. London Ser-A **370**, 173–191 (1980)
20. Hörmander, L.: An introduction to complex analysis in several variables. Amsterdam: North-Holland 1973
21. Hughston, L. P.: The twistor cohomology of local Hertz potentials. Twistor Newsletter **4**, 12–16 (1977)[13]
22. Hughston, L. P.: Some new contour integral formulae. In: Complex manifold techniques in theoretical physics (eds. D. E. Lerner, P. D. Sommers). Research Notes in Math. **32**, 115–125. San Fransisco, London, Melbourne: Pitman 1979
23. Hughston, L. P.: Twistors and particles. In: Lecture Notes in Physics 97. Berlin, Heidelberg, New York: Springer 1979
24. Lerner, D. E.: The "inverse twistor function" for positive frequency fields. Twistor Newsletter **5**, 17–18 (1977)[13]
24a. Lerner, D. E.: Twistors and induced representations of SU(2, 2). J. Math. Phys. (N.Y.) **18**, 1812–1817 (1977)
25. Malgrange, B.: Faisceaux sur des variétés analytiques-réelles. Bull. Soc. Math. France **85**, 231–237 (1957)
26. Milnor, J., Stasheff, J.: Characteristic classes. Princeton: Princeton University Press 1974
27. Penrose, R.: Zero-rest-mass fields including gravitation: asymptotic behaviour. Proc. R. Soc. London Ser. A **284**, 159–203 (1965)
28. Penrose, R.: Twistor algebra. J. Math. Phys. (N.Y.) **8**, 345–366 (1967)
29. Penrose, R.: Solutions of the zero-rest-mass equations. J. Math. Phys. (N.Y.) **10**, 38–39 (1969)
30. Penrose, R.: Twistor quantization and curved space-time. Int. J. Theor. Phys. **1**, 61–99 (1968)
31. Penrose, R.: The structure of space-time. In: Battelle rencontres 1967, pp. 121–235. New York: Benjamin 1968
32. Penrose, R., MacCallum, M. A. H.: Twistor theory: an approach to the quantization of fields and space-time. Phys. Rep. **6C**, 241–316 (1972)
33. Penrose, R.: Twistor theory, its aims and achievements. In: Quantum gravity: an Oxford Symposium (eds. C. J. Isham, R. Penrose, D. W. Sciama, pp. 268–407. Oxford: Clarendon Press 1975
34. Penrose, R.: Twistors and particles: an outline. In: Quantum theory and the structure of space-time (eds. L. Castell, M. Drieschner, C. F. von Weizsäcker,) Munich: Munich Verlag 1975
35. Penrose, R.: Non-linear gravitons and curved twistor theory. Gen. Rel. Grav. **7**, 31–52 (1976)
36. Penrose, R.: Twistor functions and sheaf cohomology. Twistor Newsletter **2**, 3–12 (1976)[13]
37. Penrose, R.: Massless fields and sheaf cohomology. Twistor Newsletter **5**, 9–13 (1977)[13]
38. Penrose, R.: The twistor programme. Rep. Mathematical Phys. **12**, 65–76 (1977)
39. Penrose, R.: Twistors as helicity raising operators. Twistor Newsletter **8**, 35–36 (1979)[13]
40. Penrose, R., Rindler, W.: Spinors and space-time structure. Cambridge: Cambridge University Press (to appear)
41. Penrose, R., Ward, R. S.: Twistors for flat and curved space-time. Einstein centennial volume (eds. P. G. Bergman, J. N. Goldberg, A. P. Held) (to appear).
42. Rawnsley, J. H.: On the Atiyah–Hitchin–Drinfeld–Manin vanishing theorem for cohomology groups of instanton bundles. Math. Ann. **241**, 43–56 (1979)
43. Serre, J-P.: Un théorème de dualité. Commun. Math. Helv. **29**, 9–26 (1955)

13 See page 45

44. Siu, Y. T.: Analytic sheaf cohomology groups of dimension n of n-dimensional noncompact complex manifolds. Pac. J. Math. **28**, 407–411 (1969)
45. Swan, R. G.: The theory of sheaves. Chicago, London: University of Chicago Press 1964
46. Ward, R. S.: The twisted photon. Twistor Newsletter 1, 2–4 (1976)[13]
47. Ward, R. S.: Curved twistor spaces. D. Phil. thesis, Oxford 1977
48. Ward, R. S.: On self-dual gauge fields. Phys. Lett. **61A**, 81–82 (1977)
49. Ward, R. S.: Sheaf cohomology and an inverse twistor function. Twistor Newsletter **6**, 13–15 (1977)[13]
50. Ward, R. S.: A class of self-dual solutions of Einstein's equations. Proc. R. Soc. London Ser. A **363**, 289–295 (1978)
51. Wells, R. O. Jr.: Differential analysis on complex manifolds. Berlin, Heidelberg, New York: Springer 1980
52. Wells, R. O. Jr.: Complex manifolds and mathematical physics. Bull. Am. Math. Soc. (new series) **1**, 296–336 (1979)
53. Wells, R. O. Jr.: Cohomology and the Penrose transform. In: Complex manifold techniques in in theoretical physics (eds. D. E. Lerner, P. D. Sommers) Research Notes in Math. **32**, 92–114. San Fransisco, London, Melbourne: Pitman 1979
54. Wells, R. O., Jr.: Hyperfunction solutions of the zero-rest-mass field equations in Commun. Math. Phys. (to appear)
55. Woodhouse, N. M. J.: Twistor cohomology without sheaves. In: Advances in twistor theory (eds. L. P. Hughston, R. S. Ward,) Research Notes in Math. 37. San Francisco, London, Melbourne: Pitman 1979

Communicated by A. Jaffe

Received April 9, 1980

13 See page 45

130
A THEORY OF 2-SURFACE ('SUPERFICIAL') TWISTORS

Twistor Newsletter **13** (1981). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

This fairly brief TN note introduces the ideas that appeared in the Royal Society Proceedings (Chapter 132) some months later.

 A theory of 2-surface ('superficial') twistors. *Twistor Newsletter* **13** (1981). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.2 **A theory of 2-surface ('superficial') twistors** *by R.Penrose* (**TN** 13, December 1981)

As had been noted some years back (Penrose 1968), the spin-lowering property of a solution in $\mathbb{M}^I$, of the symmetric 2-index twistor equation

$$\nabla^{(A}_{A'}\gamma^{BC)} = 0, \qquad \gamma^{AB} = \gamma^{BA} \tag{1}$$

can be used to reduce the expressions for the 10 conservation laws of linear gravity theory (field ϕ_{ABCD} subject to $\nabla^{AA'}\phi_{ABCD} = 0$ in the source-free region) to the charge conservation law of Maxwell theory (field $i\phi_{AB}$ subject to $\nabla^{AA'}\phi_{AB} = 0$ in a source-free space) where we set

$$\frac{G}{2}\,\phi_{AB} = \phi_{ABCD}\gamma^{CD} \tag{2}$$

Putting (in source-free region only)

$$K_{abcd} = \phi_{ABCD}\varepsilon_{A'B'}\varepsilon_{C'D'} + \bar{\phi}_{A'B'C'D'}\varepsilon_{AB}\varepsilon_{CD}$$

and

$$F_{ab} = i\phi_{AB}\varepsilon_{A'B'} - i\bar{\phi}_{A'B'}\varepsilon_{AB}\,,$$

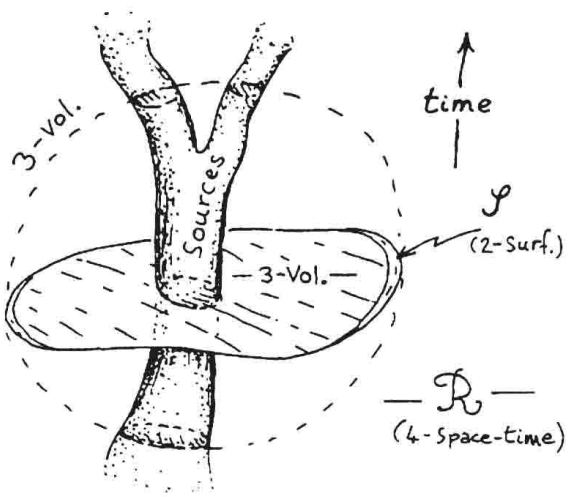
so that

$$F^{*}_{ab} = \tfrac{1}{2}e_{abcd}F^{cd} = \phi_{AB}\varepsilon_{A'B'} + \bar{\phi}_{A'B'}\varepsilon_{AB} \tag{3}$$

we can express the conservation laws for K_{abcd}, for a region $\mathfrak{R}$ surrounding a source distribution, as the fact that

$$\eta = \oint_{\mathcal{S}} F^{*}_{ab}\, dx^a \wedge dx^b \qquad (\mathcal{S} \subset \mathfrak{R}) \tag{4}$$

measures the total source for F_{ab} intercepted by any compact 3-volume whose boundary is the 2-surface $\mathcal{S}$. As γ^{AB} varies, subject to (1), η will range over the various components of the source for K_{abcd}.



Let $G^{\alpha\beta}$ be the twistor ($\in \mathbb{T}^{(\alpha\beta)}$) whose primary part is γ^{AB}. $G^{\alpha\beta}$ has 10 independent complex components, corresponding to the fact that $\mathbb{T}^{(\alpha\beta)}$ is $\frac{\boxed{4}\boxed{3}}{\boxed{2}\boxed{1}} = \frac{4\times 5}{2} = 10$-dimensional, so there are 10 independent (complex) solutions of (1). The η of (4) is *real*, so we get 20 independent real 'charges' for K_{abcd} (10 of which will be zero whenever K_{abcd} arises in the normal

way from a potential (= linearized metric)). The sources themselves are described by the various conserved currents

$$J^a = T^{ab}k_b \qquad (\nabla_a J^a = 0;\ \nabla_a T^{ab} = 0) \tag{5}$$

where k^a is a Killing vector, related to γ^{AB} by

$$k^a = \tfrac{4}{3}(i\nabla^A_{B'}\bar{\gamma}^{A'B'} - i\nabla^{A'}_{B}\gamma^{AB}). \tag{6}$$

T^{ab} is the usual energy-momentum tensor *source* for $K_{abcd}\left(K_{abcd} = K_{[cd][ab]},\ K_{[abc]d} = 0\right)$:

$$K_{acb}{}^{c} - \tfrac{1}{2}g_{ab}K_{cd}{}^{cd} = -8\pi G T_{ab} \tag{7}$$

(G = gravitational constant). In tensor terms, putting

$$S^{ab} = \gamma^{AB}\varepsilon^{A'B'} + \varepsilon^{AB}\bar{\gamma}^{A'B'} \tag{8}$$

we have

$$\nabla^{(a}S^{b)c} - \nabla^{(a}S^{c)b} + g^{a[b}\nabla_d S^{c]d} = 0 \tag{9}$$

as the translation of (1) (so equation (9) has 20 independent real solutions), and

$$e_{abcd}\nabla^b S^{cd} = \tfrac{3}{2}k_a \tag{10}$$

as the translation of (6). Equations (7), (9) and (10) imply

$$\nabla_{[a}\left(K_{bc]qd}S^{qd}\right) = \frac{4\pi G}{3}e_{abcd}T^{df}k_f \tag{11}$$

which becomes

$$\nabla_{[a}F^*_{bc]} = \frac{4\pi}{3}e_{abcd}J^d \tag{12}$$

(i.e. the Maxwell equation $\nabla_a F^{ab} = 4\pi J^b$) when

$$F^*_{ab} = \frac{1}{G}K_{abcd}S^{cd} \tag{13}$$

which reduces to (2) in source-free space.

In twistor terms, the relation between γ^{AB} and the Killing vector $k^{AB'}$(taken to be the primary part of the Hermitian twistor $K^{\alpha}{}_{\beta} \in \mathbf{T}^{\alpha}_{\beta}$) is given by

$$K^{\alpha}{}_{\beta} = 4\overline{G}_{\beta\gamma} I^{\alpha\gamma} + 4G^{\alpha\gamma} I_{\beta\gamma}. \tag{14}$$

(This is equation (6) in twistor form.) Note that the $K^{\alpha}{}_{\beta}$ of (14) is trace-free, and with this restriction $K^{\alpha}{}_{\beta}$ is determined by its primary part $k^{AB'}$. Equation (14) is part of the 'kinematic sequence' or 'moment sequence', which is (in the simplified form due to L.P.Hughston)

$$0 \to \mathbf{R} \to \mathbf{Q}^{[\alpha\beta]} \to \mathbb{H}^{\alpha}_{\beta} \to \mathbb{T}^{(\alpha\beta)} \to \mathbb{H}^{\alpha}_{\beta} \to \mathbf{Q}^{[\alpha\beta]} \to \mathbb{R} \to 0 \tag{15}$$

where these are considered as *real* vector spaces, of respective dimensions 0, 1, 6, 15, 20, 15, 6, 1, 0, $\mathbf{H}^{\alpha}_{\beta}$ consisting of trace-free Hermitian twistors H^{α}_{β} and $\mathbf{Q}^{[\alpha\beta]}$ consisting of twistor-real (i.e. $\overline{Q}_{\alpha\beta} = \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta} Q^{\gamma\delta}$) elements $Q^{\alpha\beta} \in \mathbb{T}^{[\alpha\beta]}$. Equation (14) is the $\mathbf{T}^{(\alpha\beta)} \to \mathbf{H}^{\alpha}_{\beta}$ part of (15), so twistors whose primary parts are Killing vectors constitute the image of this map. The kernel provides those γ^{AB}'s which give zero in (6) (i.e. in (10)). This kernel is the image of $\mathbf{H}^{\alpha}_{\beta} \to \mathbb{T}^{(\alpha\beta)}$, which is the set of twistors of the form

$$iH^{(\alpha}_{\gamma} I^{\beta)\gamma} \tag{16}$$

The η's of (4), with (13) substituted, arise as real parts of expressions like

$$\zeta = \frac{1}{16\pi G} \oint_{\mathcal{S}} K_{AA'BB'cd} \gamma^{AB} \varepsilon^{A'B'} dx^{c} \wedge dx^{d} \tag{17}$$

which we can write

$$\zeta = A_{\alpha\beta} G^{\alpha\beta} \tag{18}$$

where $A_{\alpha\beta}$ is the total *kinematic* or *angular momentum twistor* of the sources surrounded by $\mathcal{S}$ and satisfies

$$A_{\alpha\beta} I^{\gamma\beta} = \overline{A}^{\gamma\beta} I_{\alpha\beta}. \tag{19}$$

$A_{\alpha\beta}$ is dual (over the reals) to $G^{\alpha\beta}$, so it is located in the dual sequence to (15), which is another sequence of just the same form as (15), $A_{\alpha\beta}$ lying in the image of the corresponding map $\mathbf{H}^{\alpha}_{\beta} \rightarrow \mathbf{T}_{(\alpha\beta)}$, which is what (19) entails.

All this refers just to $\mathbf{M}^I$; we wish to do a corresponding thing in a general curved space $\mathcal{M}$. We cannot expect to solve (1) (or (9)) in general, but we need its appropriate analogue *only on the surface* $\mathcal{S}$ *itself.* However, if we try to restrict (1) to $\mathcal{S}$, we find only *two* equations that are entirely tangential to $\mathcal{S}$, which is insufficient to determine the *three* complex components of γ^{AB} on $\mathcal{S}$. The key idea is not to try to define the γ's directly but to consider them as tensor products of 1-index objects. Any solution of (1) is, indeed, of the form

$$\gamma^{AB} = \sum \underset{1}{\omega}{}^{(A} \underset{2}{\omega}{}^{B)} \tag{20}$$

in $\mathbf{M}^I$, where each $\underset{i}{\omega}{}^{A}$ satisfies the 1-index twistor equation

$$\nabla^{(A}_{A'} \omega^{B)} = 0 \tag{21}$$

and it is (21) rather than (1) that we restrict to $\mathcal{S}$. We take $\mathcal{S}$ to be *spacelike* and (normally) to have the *topology of a sphere* S^2. Using the GHP formalism (Geroch, Held & Penrose 1973) we find, for the two components tangential to $\mathcal{S}$ (in a spin-frame family whose flagpoles are orthogonal to $\mathcal{S}$), that

$$\eth' \omega^0 = \sigma' \omega^1, \quad \eth \omega^1 = \sigma \omega^0. \tag{22}$$

It follows from the Atiyah-Singer index theorem that equations (22) always have at least 4 (complex-) independent solutions and that *generically* equations (22) have *precisely* 4 independent solutions. (In the 'canonical' case, $\sigma = \sigma' = 0$, and the adjoint equations have no solutions.) The solution space $\mathbf{T}^{\alpha}(\mathcal{S})$ is the *2-surface twistor space* of $\mathcal{S}$, which will be a complex 4-dimensional vector space, assuming that we are in this normal situation. Then the symmetric tensor product $\mathbb{T}^{(\alpha\beta)}(\mathcal{S})$ will be 10-complex-dimensional. Each element of $\mathbb{T}^{(\alpha\beta)}(\mathcal{S})$ can be interpreted as a spinor field γ^{AB} over $\mathcal{S}$ (the tensor product of the ω^A's being taken pointwise on $\mathcal{S}$.) The quantities (17) can now be defined in a *general* $\mathcal{M}$, as before but with Riemann's R_{abcd} replacing K_{abcd}.

For a 1-term sum (20) we get

$$\zeta = \frac{-i}{4\pi G}\oint_{\mathcal{S}} \{(\Psi_1 - \Phi_{10})\omega_1^0\omega_2^0 + (\Psi_2 - \Phi_{11} - \Lambda)(\omega_1^0\omega_2^1 + \omega_1^1\omega_2^0) + (\Psi_3 - \Phi_{21})\omega_1^1\omega_2^1\}\underset{\sim}{\mathcal{S}} \qquad (23)$$

($\underset{\sim}{\mathcal{S}}$ being the surface-area 2-form on $\mathcal{S}$). This defines $A_{\alpha\beta} \in \mathbf{T}_{(\alpha\beta)}(\mathcal{S})$ via (18), which has 10 *complex* components. To define mass-momentum and angular momentum *quasi-locally* we need an analogue of (19) so as to have only 10 *real* components. So far this is lacking; we do not have a convincing $I_{\alpha\beta} \in \mathbf{T}_{[\alpha\beta]}(\mathcal{S})$, nor a convincing $Z^\alpha \mapsto \overline{Z}_\alpha$ (through tentative definitions exist for the latter—not conformally invariant whereas $\mathbf{T}^\alpha(\mathcal{S})$ is). Things are rosier at $\mathcal{I}^+$, where $\sigma' = 0$. $I_{\alpha\beta}$ exists via

$$I_{\alpha\beta}Z_1^\alpha Z_2^\beta = i\omega_1^0\pi_{21'} - i\omega_2^0\pi_{11'} = \omega_2^0\eth\omega_1^0 - \omega_1^0\eth\omega_2^0 \qquad (24)$$

being constant over $\mathcal{S} \subset \mathcal{I}^+$, where

$$\pi_{0'} = i\eth'\omega^1 - i\rho\omega^0,\ \pi_{1'} = i\eth\omega^0 - i\rho'\omega^1. \qquad (25)$$

Still $Z^\alpha \mapsto \overline{Z}_\alpha$ is not totally convincing, but the *Bondi-Sachs* 4-momentum is reproduced exactly and a *new improved angular momentum at* $\mathcal{I}^+$ is obtained.

Note K.P.Tod's tantalizing formula for (23), using (25), $\zeta = \frac{-i}{4\pi G}\oint_{\mathcal{S}} (\pi_{0'}\pi_{1'} + \pi_{1'}\pi_{0'})\underset{\sim}{\mathcal{S}}$.

More work in progress; see Penrose (1982). Thanks especially to I.M.Singer and K.P.Tod.

References

Geroch, Held & Penrose (1973) J. Math Phys, **14**, 874.

Penrose, R. (1968) Int. J. Theor. Phys. **1**, 61.

Penrose, R. (1982) Proc. Roy. Soc. Lond. **A381**, 53.

131
SOME UNSOLVED PROBLEMS IN CLASSICAL GENERAL RELATIVITY

Seminar on differential geometry, 631–68. *Ann. of Math. Stud.*, 102, Princeton University Press, 1982

162 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

SOME UNSOLVED PROBLEMS IN CLASSICAL GENERAL RELATIVITY

R. Penrose

Although a great deal of the research into general relativity carried out at present is concerned with its relation to quantum theory, there are many interesting outstanding problems in the original classical theory which have remained unsolved for many years. In view of the important recent work of Schoen and Yau (1979, 1980), which establishes the positive-definiteness of mass in general relativity when measured at spatial infinity, it would seem reasonable to hope that others of these classical problems of general relativity may also be resolved in the not-too-distant future. It should be pointed out, moreover, that general relativity has become, in recent years, an experimentally well-tested theory. So any mathematical results of importance to the classical theory will be assured a permanent place in physics. But the same cannot at all be said of results relating to current work in the quantum theory of gravity, since that remains a highly speculative collection of physical and mathematical ideas which cannot yet be properly said to constitute a theory at all, let alone a well-tested one.

I shall make no attempt here to be at all comprehensive in the classical relativity problems I propose to discuss. They are mainly related to the lines of work that I have myself chosen to pursue. There are undoubtedly many other unresolved problems of equal interest to others.

The conventions adopted here are: the metric tensor g_{ab} (and its inverse g^{ab}) have signature $(+,-,-,-)$; lower case latin letters denote space-time indices; the symbol ∇ (or ∇_a) stands for (Christoffel) covariant derivative; Riemann tensor sign conventions are fixed by $(\nabla_a\nabla_b - \nabla_b\nabla_a)q^d = R_{abc}{}^d q^c$, with $R_{ab} = R_{acb}{}^c$; physical units are taken so that $G = c = 1$.

The result of Schoen and Yau refers to space-times which are appropriately asymptotically flat at spatial infinity. Asymptotically flat space-times are interesting, not because they are thought to be realistic models for the entire universe, but because they describe the gravitational fields of isolated systems, and because it is only with asymptotic flatness that general relativity begins to relate in a clear way to many of the important aspects of the rest of physics, such as energy, momentum, radiation, etc. Asymptotic flatness is thus a highly significant physical idealization, and it is gratifying that certain apparently appropriate conditions for a space-time to be asymptotically flat can be put in an elegant geometrical form, —though exactly *how* appropriate these conditions are is itself an unresolved issue.

In detail, the asymptotic conditions that one chooses to impose may depend on the nature of the problem at hand. For the Schoen-Yau theorem, suitable conditions need to be placed at spatial infinity, and this is adequate for the discussion of total mass-energy of a system, where one is not concerned with questions of radiation; such as mass carried away from a system by gravitational waves. To discuss radiation problems it is more appropriate to impose conditions at *null* infinity.

A glance at Figure 1 should explain why this is so. The intuitive meaning of the mass defined at spatial infinity is that it is the total integrated mass-energy that is intercepted by some spacelike hypersurface approaching a flat spacelike hyperplane at infinity. The mass carried away by outgoing radiation (whether electromagnetic, gravitational, or of any other massless field) travels outwards from the source essentially with the speed of light, and this would always be incorporated in the total

Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

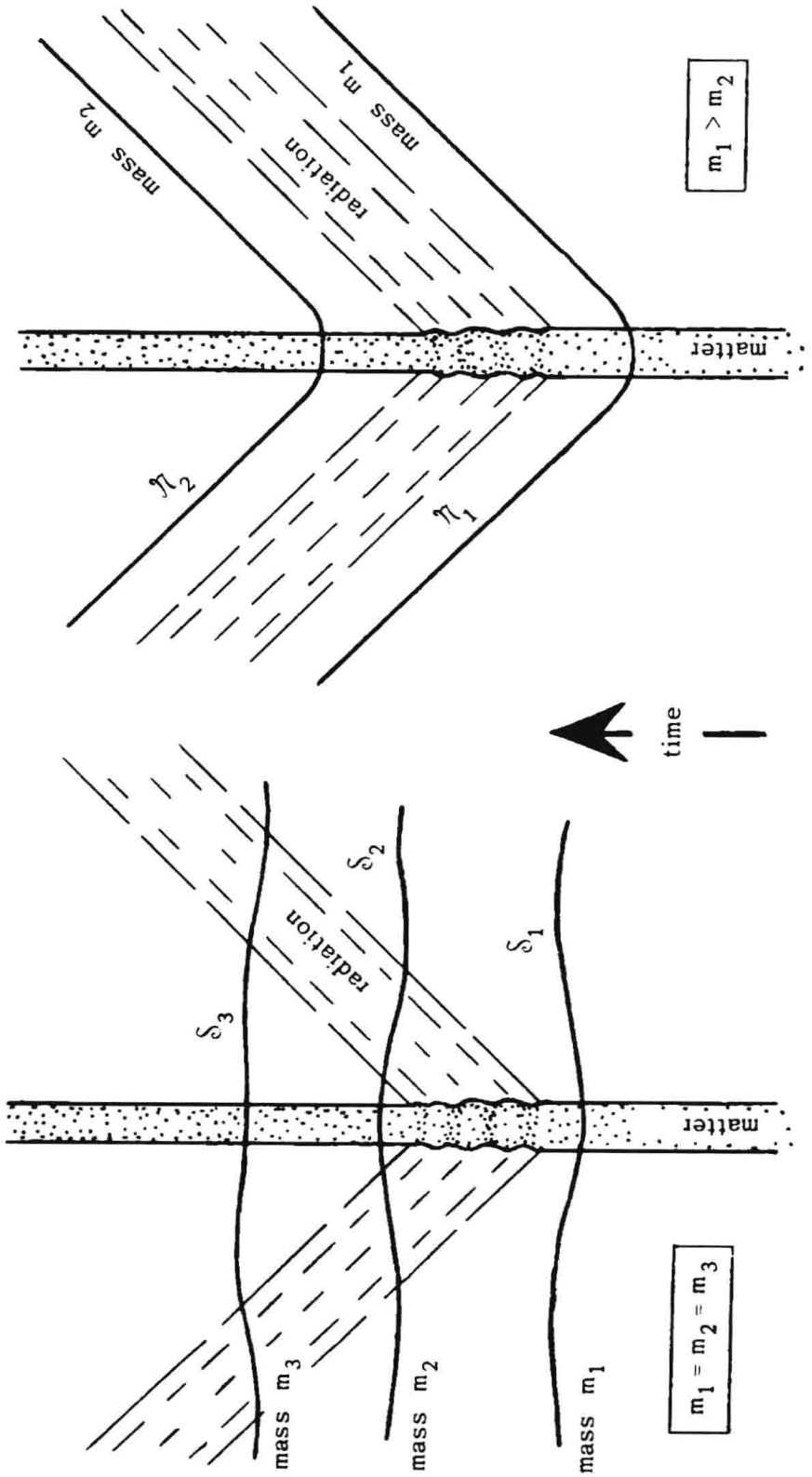


Figure 1

(i) the total mass-energy intercepted by spacelike hypersurfaces $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ remains constant and one cannot ascertain the portion of the mass carried away by radiation.

(ii) by using (asymptotically) *null* hypersurfaces $\mathcal{H}_1$, $\mathcal{H}_2$ we can infer the mass-loss due to radiation.

mass measure no matter how far into the future the spacelike hypersurface is chosen. If the *total* mass-energy is the only quantity that we have, then its measure on each spacelike hypersurface will stay the same and will not enable us to distinguish the part residing in the radiation from that part which remains in the source. However if a *null* or asymptotically null hypersurface is used, which opens out into the future, then radiation can escape between two such hypersurfaces, the difference between the total mass intercepted by the earlier hypersurface and that intercepted by the later one being a measure of the mass carried by the escaping radiation.

It should be pointed out that the reason we are forced into considerations of this kind is that there is no *local* definition of mass-energy in general relativity which takes into account the contribution coming from the gravitational field itself. (That the gravitational field itself must have mass follows from the fact that the potential energy of a pair of gravitating particles depends upon their distance apart, so that *total* energy, and therefore the total mass, of the system particle+particle+field must depend upon this distance and therefore on the gravitational field configuration.) To appreciate the essentially non-local nature of gravitational energy, we recall that in *special* relativity we have, in the flat Minkowski space-time background, a local energy-momentum tensor T_{ab} which is symmetric and which satisfies a vanishing divergence law

$$\nabla^a T_{ab} = 0 . \tag{1}$$

Taking the four components of this equation with respect to a standard constant basis frame, we have four independent equations of the type

$$\nabla^a J_a = 0 , \tag{2}$$

where J_a is each of T_{a0}, T_{a1}, T_{a2}, T_{a3} in turn.

Equation (2) is, of course, simply $dJ^* = 0$ where the 3-form J^* is the Hodge dual of the covector J, so J provides us with a conserved current. In the case of electromagnetic theory we have one such J,

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

providing the familiar charge conservation law: the total electric charge entering a compact space-time 4-volume equals the total charge leaving it—and the total charge intercepting any hypersurface is measured simply by the integral of J^* over it. Here we have four such conservation laws, providing the four components of the conserved energy-momentum 4-vector.

One of the cornerstones of *general* relativity is that the equation (1) should remain true in that theory. But since we do not now have constant basis frames, we cannot obtain a globally conserved energy-momentum simply by integrating components of T_{ab}. From the physical point of view this is reasonable. For in Einstein's equation

$$R_{ab} - \frac{1}{2} R g_{ab} = -8\pi T_{ab} \tag{3}$$

(where, usually, the gravitational constant G is inserted after "8π"; but we have adopted units with $G = 1$), the T_{ab} on the right-hand side satisfies (1) and describes the energy-momentum of all the matter fields (i.e. everything except gravity itself). The contribution to the energy-momentum from the *gravitational* field does not appear in T_{ab}, so it is not to be expected that T_{ab} alone should provide us with a conserved energy-momentum. Instead, the gravitational contribution enters non-locally, showing up in the total energy-momentum as measured at infinity but not showing up in the components of any locally defined tensor quantity.

It may be, however, that some sort of *quasi-local* definition of energy-momentum in general relativity exists, where one does not need to go "all the way to infinity" in order for the concept to be meaningfully defined, but where, for example, a total energy-momentum might be assigned to any 2-surface, $\partial\mathcal{H}$ bounding a compact portion $\mathcal{H}$ of a spacelike hypersurface, the physical interpretation of this energy-momentum being the total energy-momentum (gravitational plus that due to matter fields) intercepted by $\mathcal{H}$. To be useful, such a quasi-local energy-momentum concept should be subject at least to certain inequalities expressing positivity and semi-additivity requirements. Thus, we have, for our first problem:

PROBLEM 1. Find a suitable quasi-local definition of energy-momentum in general relativity.

In special relativity, such a quasi-local definition exists, as we have seen, the gravitational contribution being zero, and the answer being obtained simply by integrating the appropriate components of T_{ab} over $\mathcal{H}$. Another way of thinking of these components is that they are given by currents of the form

$$J_a = T_{ab}k^b$$

where k is a Killing vector generating a translational symmetry of the Minkowski space-time. But J still satisfies (2) when k is any Killing vector whatever, such as that generating a rotational motion of the space-time. In this way we obtain *ten* conserved currents rather than just four, the six new ones providing the components of the total relativistic angular momentum (including those describing the motion of the mass-centre). Thus, there is a more ambitious successor to Problem 1:

PROBLEM 2. Find a suitable quasi-local definition of angular momentum in general relativity.

To expect a positive solution to Problems 1 and 2 may well be felt to be excessively optimistic. But the vagueness of their statements allows for a certain latitude in their interpretations. In fact a suggested approach to these problems was provided some years ago, namely the "linkages" of Tamburino and Winicour (1966, cf. Winicour 1980), but the resulting definitions involve some ambiguities and lack adequate useful inequalities. In the absence of satisfactory solutions to Problems 1 and 2, therefore, we appear to be forced into the asymptotic considerations that we opened with.

One of the important landmarks of general relativity in the post-Einstein period was the work of Bondi (1960, Bondi *et al.* 1962) and its generalization due to Sachs (1962a) which showed that, with a suitable

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

definition at null infinity, the mass carried by gravitational waves is necessarily positive-definite. Thus, in Figure 1 we have, with the Bondi-Sachs definition, $m_1 \geq m_2$ provided that $\mathcal{I}_2$ lies entirely to the future of $\mathcal{I}_1$. For this result, it is required that an appropriate condition of approach to flatness at infinity is imposed in future-null directions. In the initial work of Bondi and Sachs this was done in terms of the existence of a certain type of coordinate system in which a certain coordinate u called "retarded time" takes constant values on outgoing null hypersurfaces (this being valid outside a certain spatially compact region—or world-tube—which contains all the matter). Thus $\mathcal{I}_1$ could be given by $u = u_1$ and $\mathcal{I}_2$ by $u = u_2$. In fact the coordinate conditions imposed by Bondi and Sachs imply a certain restriction on the relation between the hypersurfaces $\mathcal{I}_1$ and $\mathcal{I}_2$, namely that one should be obtained from the other by a so-called "translation" rather than by a more general "supertranslation," but this restriction is actually not necessary for the mass-loss inequality $m_1 \geq m_2$ (Penrose (1964), (1967)). The meaning of the terms "translation" and "supertranslation" will be given shortly. These concepts are best understood in terms of a formalism somewhat different from that of Bondi and Sachs, and this will be described next.

The idea is to borrow from projective geometry or, more appropriately from complex function theory, the concept of "points at infinity." We envisage a conformal factor Ω which rescales the space-time metric g_{ab} to a new one

$$\hat{g}_{ab} = \Omega^2 g_{ab} ,$$

the scalar field Ω approaching zero, at space-time infinity, at such a rate that $\hat{g}_{ab}$ smoothly attains finite values there. The possibility of introducing such a conformal factor is, in effect, equivalent to the Bondi-Sachs asymptotic flatness condition.

The fix idea a little more clearly, let us see how this works in the case of Minkowski space. The standard form of the metric is

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2$$

but it is convenient first to choose new coordinates u, v, θ, ϕ where

$$u = t - r, \quad v = t + r, \quad \text{with} \quad r = (x^2+y^2+z^2)^{1/2},$$

and where

$$z = r\cos\theta, \quad x = r\sin\theta\cos\phi, \quad y = \sin\theta\sin\phi$$

so the metric form becomes

$$ds^2 = dudv - \frac{1}{4}(v-u)^2(d\theta^2 + \sin^2\theta d\phi^2).$$

The coordinate u is, in fact, an example of a Bondi-Sachs retarded time parameter (and v is a Bondi-Sachs advanced time parameter), so $u =$ const. would provide examples of the outgoing null hypersurfaces referred to earlier. Any suitable conformal factor Ω has to have, as it turns out, an asymptotic behavior that is like the reciprocal of an affine parameter on null geodesics. Thus, along each null line $u, \theta, \phi =$ const. we expect $\Omega \sim v^{-1}$ in future directions, and similarly $\Omega \sim u^{-1}$ in past directions, along $v, \theta, \phi =$ const. This suggests the choice

$$\Omega = \{(1+u^2)(1+v^2)\}^{-1/2}$$

(although there is much arbitrariness in Ω).

The further coordinate change to

$$p = \tan^{-1}v, \quad q = \tan^{-1}u$$

provides us with a regular coordinate system p, q, θ, ϕ "at infinity" (where $\Omega = 0$), the rescaled metric $d\hat{s} = \Omega ds$ being now given by

$$ds^2 = dpdq - \frac{1}{4}\sin^2(p-q)(d\theta^2 + \sin^2\theta d\phi^2).$$

The original Minkowski space corresponds to the range $-\frac{1}{2}\pi < q \leq p < \frac{1}{2}\pi$

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

(where $\Omega > 0$), but we can widen this to include the boundary points, at which $q = -\frac{1}{2}\pi$ or $p = \frac{1}{2}\pi$. The relevant coordinate range is indicated in Figure 2.

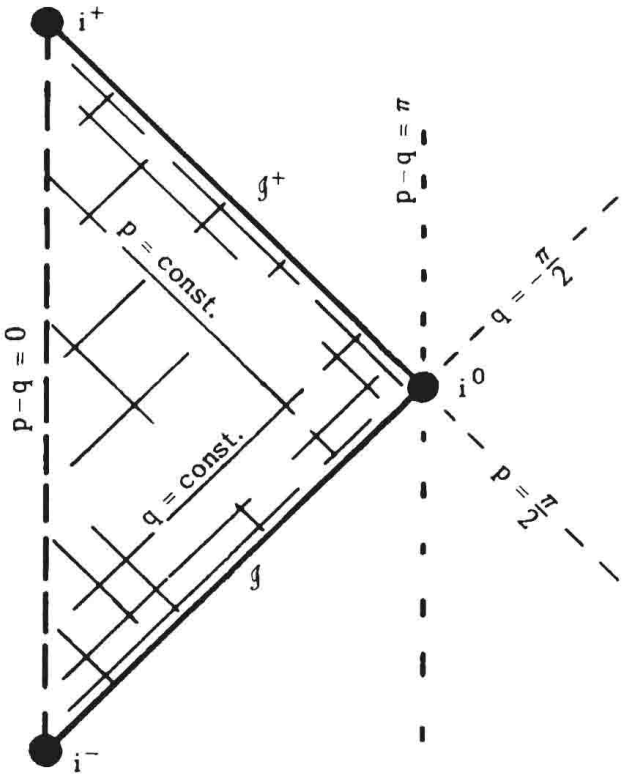


Figure 2

The interior of the triangular region (together with the left-hand symmetry axis) traces out the region of the Einstein universe that is conformal to Minkowski space.

Each point of the triangular (p,q)-region describes a 2-sphere whose radius (in the $d\hat{s}$ metric) is $\frac{1}{2}\sin(p-q)$ —except that each point of the "symmetry axes" $p = q$ and $p - q = \pi$ represents a single point. If we extend the coordinate range of p, q to the whole of that lying between these axes (i.e. to the range $0 \leq p - q \leq \pi$), then we obtain the metric of the "Einstein universe," which is topologically $S^3 \times \mathbf{R}$. (This is a static model which is spatially an exact 3-sphere.) The region lying between the null hypersurfaces $\mathcal{I}^-$ (given by $q = -\frac{1}{2}\pi$, $-\frac{1}{2}\pi < p < \frac{1}{2}\pi$)

and $\mathscr{I}^+$ (given by $p = \frac{1}{2}\pi$, $-\frac{1}{2}\pi < q < \frac{1}{2}\pi$) is the part conformal to Minkowski space, but the ds metric is regular (with $\Omega = 0$) also at all points of $\mathscr{I}^-$ and $\mathscr{I}^+$. We refer to $\mathscr{I}^-$ as past null infinity and $\mathscr{I}^+$ as future null infinity for Minkowski space. Any null geodesic acquires a past end-point on $\mathscr{I}^-$ and a future end-point on $\mathscr{I}^+$. (We recall that the concept of a null geodesic is conformally invariant—it is the same with respect to g_{ab} as it is with respect to $\hat{g}_{ab}$.)

The well-known Schwarzschild and Kerr metrics (cf. Hawking and Ellis (1973) are other examples for which conformally regular boundary hypersurfaces $\mathscr{I}^-$ and $\mathscr{I}^+$ can be introduced. These metrics satisfy Einstein's vacuum equations (i.e. are Ricci-flat), but they possess singularities in the interior regions. ("Singularities" may be recognized in terms of geodesic incompleteness, if desired.) The singularities may be removed in the interior by replacing the metric there by another one in which the vacuum equations are violated (i.e. $R_{ab} \neq 0$), and the resulting Ricci tensor is interpreted, according to Einstein's equations, as referring to a mass-energy (matter) region which acts as a source for the external gravitational field. The space-time with its boundary $\mathscr{I}^- \cup \mathscr{I}^+$ has a structure like that depicted in Figure 3 (with one spatial dimension being suppressed). Again every null geodesic acquires an end-point on each of $\mathscr{I}^-$ and $\mathscr{I}^+$. The manifold-with-boundary $\overline{\mathscr{M}}$ that we obtain (and whose interior is conformal to the physical space-time $\mathscr{M}$) is not quite compact. Roughly speaking, it has "three points missing." These would be the points $i^-(p = q = -\frac{1}{2}\pi)$, $i^0(p = -q = \frac{1}{2}\pi)$ and $i^+(p = q = \frac{1}{2}\pi)$ in the Minkowski case (representing past-timelike, spacelike and future-timelike infinities, respectively), but they would generally be singular points for the conformal metric of $\overline{\mathscr{M}}$, when $\mathscr{M}$ is asymptotically flat but not flat, and so must be excluded.

This discussion motivates the definition of a space-time $\mathscr{M}$, with metric g, as being k-*asymptotically simple* (Penrose 1965b) if there exists a manifold-with-boundary $\overline{\mathscr{M}}$, with C^k metric g ($k \geq 2$) and boundary $\mathscr{I}$, such that

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

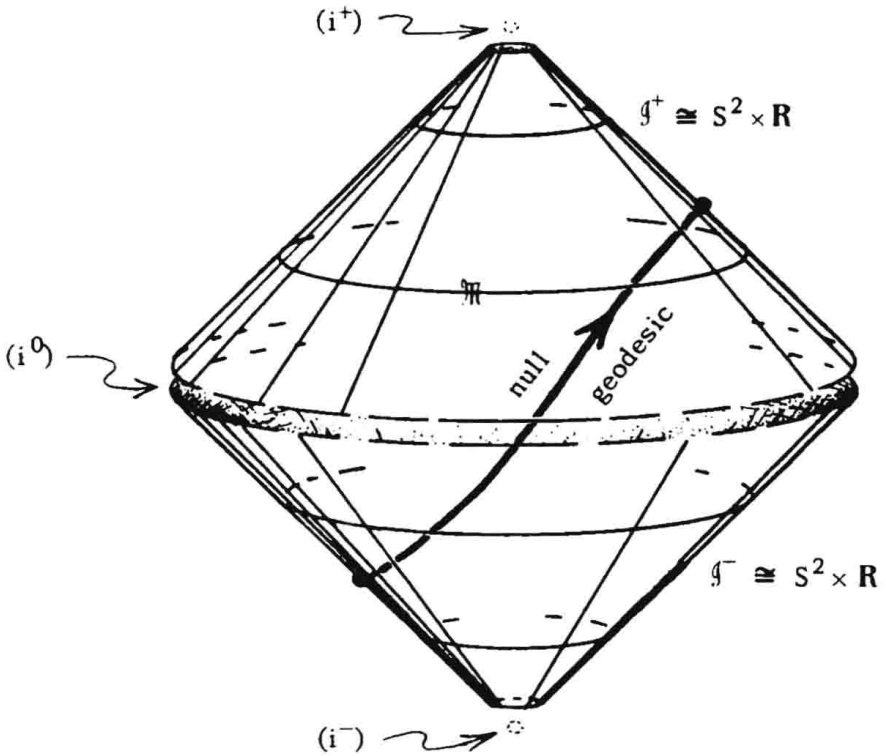


Figure 3

$\overline{\mathscr{M}}$ for asymptotically simple space-time $\mathscr{M}$.

(i) $\overline{\mathscr{M}} - \mathscr{I}$ is conformal to $\mathscr{M}$, with $\hat{g} = \Omega^2 g$,

(ii) Ω is C^k throughout $\overline{\mathscr{M}}$, while $\Omega = 0$ and $d\Omega \neq 0$ on $\mathscr{I}$,

(iii) every null geodesic in $\mathscr{M}$ acquires two end-points on $\mathscr{I}$.

If Einstein's vacuum equations hold for g in the neighborhood of $\mathscr{I}$, then it follows that $\mathscr{I}$ is *null* and consists of two disjoint pieces $\mathscr{I}^-$ and $\mathscr{I}^+$ each having topology $S^2 \times \mathbf{R}$, just as in the Minkowski, Schwarzschild and Kerr cases (Penrose 1965b, Geroch 1971). Furthermore $\mathscr{M} \cong \mathbf{R}^4$ (Geroch 1971).

Asymptotic simplicity has the appearance of being a very reasonable condition to impose on a space-time which describes an isolated system in general relativity and in which all light rays escape to infinity. And one would anticipate that there are many asymptotically flat Ricci-flat

space-times representing imploding-exploding gravitational waves (where the intensity and localization is insufficient to produce collapse to a black hole). Yet the following is at present an unsolved problem:

PROBLEM 3. Find an asymptotically simple Ricci-flat space-time which is not flat—or at least prove that such space-times exist.

A related problem is:

PROBLEM 4. Are there restrictions on k for non-stationary k-asymptotically simple space-times, with non-zero mass, which are vacuum near $\mathscr{I}$?

(A stationary space-time is one possessing a timelike Killing vector field; "vacuum near $\mathscr{I}$" means that the g-metric is Ricci-flat in some neighborhood of $\mathscr{I}$ in $\overline{\mathscr{M}}$.)

Some work by Schmidt and Stewart (1979) and by Porrill and Stewart (1980) rather suggests that an affirmative answer to Problem 4 may perhaps be the case, with k having a rather small value. If this turns out to be so it would make the asymptotic analysis of isolated gravitating systems much less geometrically pleasant than one had originally hoped. The difficulty seems to be a possible incompatibility between a high degree of differentiability on $\mathscr{I}^-$ and on $\mathscr{I}^+$ whenever the total mass is non-zero. In effect, i^0 is a singular point off which the gravitational (curvature) field can "scatter," possibly leading to a loss of differentiability on $\mathscr{I}^+$ or $\mathscr{I}^-$. However, the calculations which have been carried out so far depend on a linear perturbation analysis and it is not really clear what the rigorous results should actually be.

One consequence of Einstein's vacuum equations holding (for g) near $\mathscr{I}$ is that the generators of $\mathscr{I}$ (i.e. the null geodesic curves lying on $\mathscr{I}^{\pm}$, these curves constituting a fibration of $\mathscr{I}^{\pm}$) are free of shear (Penrose 1965b). This has the effect that any two spacelike cross-sections of $\mathscr{I}^+$ (or $\mathscr{I}^-$) are mapped *conformally* to one another by these generators. Since these cross-sections are all topologically 2-spheres, it follows that the

induced $\hat{g}$-metric on each can, with suitable choice of Ω, be chosen to be (minus) that of a unit sphere, so the metric on $\mathscr{I}^+$ (and similarly $\mathscr{I}^-$) becomes

$$-d\hat{s}^2 = d\theta^2 + \sin^2\theta d\phi^2 + 0 \cdot du^2 . \tag{4}$$

The final term is included as a reminder that $\mathscr{I}^+$ is *three*-dimensional, the third coordinate u not actually contributing to the form of the metric. The $d\hat{s}$ metric is in fact degenerate (vanishing determinant) with signature $(-, -, 0)$.

The conformal $d\hat{s}$ metric (4) is considered to be part of the universal intrinsic structure of $\mathscr{I}^+$ (universal, in the sense that *any* space-time which is asymptotically simple and vacuum near $\mathscr{I}$ has a $\mathscr{I}^+$ metric—and similarly a $\mathscr{I}^-$ metric—which is conformal to (4)). In fact there is a somewhat stronger (universal) intrinsic structure (Penrose 1964, 1974) that can be assigned to $\mathscr{I}^\pm$, sometimes referred to as its *strong conformal geometry*. One way of describing this structure is to say that in addition to the conformal $d\hat{s}$ metric (4), the ratio

$$d\hat{s} : du \tag{5}$$

is taken as invariant. Another way to phrase this is to say that a tensor of the form

$$\hat{g}_{\alpha\beta}\, \hat{n}^\gamma\, \hat{n}^\delta \tag{6}$$

is specified on $\mathscr{I}^\pm$, where Greek indices refer to intrinsic directions within $\mathscr{I}^\pm$, $\hat{g}_{\alpha\beta}$ being the metric tensor defined by $d\hat{s}^2$, and $\hat{n}^\gamma$ pointing along the generators of $\mathscr{I}^\pm$ (in the future direction) with $\hat{n} = \partial/\partial u$. We have

$$\hat{n}^\alpha\, \hat{g}_{\alpha\beta} = 0$$

expressing the relation between $\hat{n}$ and the degenerate direction of $\hat{g}_{\alpha\beta}$. The particular strong conformal geometry required for $\mathscr{I}^\pm$ is determined by the condition that, on $\mathscr{I}^\pm$,

$$\mp\, \hat{n}^a = \hat{g}^{ab}\, \hat{\nabla}_b\, \Omega . \tag{7}$$

An important condition on $\mathcal{I}^{\pm}$ is that its generators be *infinitely long* in the sense that the parameter u, as defined above in relation to the structure (5), (6), should cover the entire range $(-\infty, \infty)$ for each θ, ϕ (Geroch and Horowitz 1978). This suggests:

PROBLEM 5. Find conditions on the Ricci tensor R_{ab} throughout $\mathcal{M}$ which ensure that the generators of $\mathcal{I}$ are infinitely long.

A possible such condition might be the *null convergence condition*

$$R_{ab}\ell^a\ell^b \leq 0 \quad \text{whenever} \quad \ell^a\ell_a = 0 \tag{8}$$

which, with Einstein's equation (3), is the physically reasonable *weak energy condition*

$$T_{ab}\ell^a\ell^b > 0 \quad \text{whenever} \quad \ell^a\ell_a = 0 .$$

(Sign conventions have been chosen so that (8) indeed expresses convergence rather than divergence.) With asymptotic simplicity holding, (8) implies that every null geodesic in $\mathcal{M}$ either possesses a pair of conjugate points or else only just fails to do so in the sense that its end-points on $\mathcal{I}^+$ and $\mathcal{I}^-$ are conjugate. To say that the generators of $\mathcal{I}$ are infinitely long is to say that these generators also only just fail to possess pairs of conjugate points in an appropriate corresponding sense.

The generators of $\mathcal{I}^+$ being assumed infinitely long, we obtain the Bondi-Metzner-Sachs group (BMS group) as the intrinsic symmetry group of either $\mathcal{I}^+$ or $\mathcal{I}^-$, in which the strong conformal geometry (and orientations) are preserved. The BMS group consists of transformations of the form

$$\theta \mapsto \Theta(\theta, \phi)$$
$$\phi \mapsto \Phi(\theta, \phi)$$
$$u \mapsto K(\theta, \phi)\{u + A(\theta, \phi)\}$$

where

$$d\hat{s}^2 \mapsto \{K(\theta, \phi)\}^2 d\hat{s}^2$$

the functions Θ and Φ defining a (non-reflective) conformal motion of S^2 (so the function K is restricted in having a specific form, with three free real parameters) and $A(\theta, \phi)$ being an *arbitrary* (appropriately smooth) real-valued function on S^2.

The subgroup of the BMS group for which the θ, ϕ transformation is the identity on S^2 is the group of *supertranslations* of $\mathcal{I}^{\pm}$. A general supertranslation is given by $\theta \mapsto \theta$, $\phi \mapsto \phi$, $u \mapsto u + A(\theta, \phi)$ where A is arbitrary. The particular cases given when A is composed only of spherical harmonics of zeroth and first order are called *translations*. The reason for this name is that in the case of Minkowski space, any translation in the ordinary sense does in fact induce a translation on $\mathcal{I}^{\pm}$ in the sense just described. In fact, the translation subgroup of the BMS group is uniquely singled out by its group-theoretic properties, as the only 4-parameter normal subgroup of the BMS group (Sachs 1962b).

In the case of Minkowski space, with the specific u-parameter that was introduced earlier, we may identify the restricted* Lorentz group as the subgroup of the (restricted) BMS group $\mathbf{B}$ for which $A(\theta, \phi) \equiv 0$, i.e. for which the cross-section $u = 0$ of $\mathcal{I}^{+}$ is left invariant. In fact $u = 0$ is the future light cone of the (t, x, y, z)-origin O, which is invariant under standard Lorentz Transformations. The restricted Poincaré (i.e. inhomogeneous Lorentz) group is generated by these transformations together with the translations—and so may be identified as a subgroup $\mathbf{P}$ of $\mathbf{B}$. There are however, many subgroups of $\mathbf{B}$ which are isomorphic with the restricted Poincaré group. Specifically, if we take any supertranslation S and conjugate $\mathbf{P}$ with respect to it we get a distinct such subgroup $\mathbf{P}' = S\mathbf{P}S^{-1}$. As far as the group structure of $\mathbf{B}$ is concerned—and indeed, as far as the geometric structure of $\mathcal{I}^{+}$ is concerned, $\mathbf{P}$ and $\mathbf{P}'$ are completely equivalent to one another. Indeed, it is one of the characteristic difficulties of gravitational radiation theory that one must come to terms with this fact. The Poincaré group is the symmetry group of flat space-time,

*Here the "restricted ..." is to be interpreted as "connected component of the ...."

so it might have been thought that a suitably *asymptotically* flat space-time should, in some appropriate sense, have the Poincaré group as an asymptotic symmetry group. Instead, it turns out that in general we seem only to obtain the BMS group (which has the unpleasant feature of being an infinite-parameter group) as the asymptotic symmetry group of an asymptotically flat space-time.

The group $\mathbf{P}$, for Minkowski space $\mathcal{M}$, is the subgroup of $\mathbf{B}$ which leaves invariant not only the strong conformal geometry of $\mathcal{I}^+$ but also a family of cross-sections of $\mathcal{I}^+$ referred to as *good cuts*. The good cuts are characterized by the fact that they are the intersections with $\mathcal{I}^+$ of *future light cones* of points of $\mathcal{M}$. They are obtained from the particular good cut $u = 0$ by translations, and their equations are given, generally, by setting u equal to a function of θ and ϕ which is composed only of zeroth and first order spherical harmonics. The difficulty, in the case of a general asymptotically simple $\mathcal{M}$, lies in the fact that there seems to be no suitable family of cross-sections of $\mathcal{I}^+$ that can properly take over the role of the Minkowskian good cuts. This means that although the translation elements of $\mathbf{B}$ are canonically singled out, there is no canonical concept of a "supertranslation-free Lorentz rotation." These Lorentz rotations, in the Minkowski case, are distinguished as BMS transformations leaving some good cut invariant.

The physical concepts of energy and momentum are associated with the translations of Minkowski space, while the concept of angular momentum is associated with Lorentz rotations of Minkowski space. (Recall the discussion of the role of a Killing vector k in the discussion preceding Problem 2.) In accordance with this, it turns out that for an asymptotically flat space-time a good concept of energy-momentum exists at $\mathcal{I}$, namely the Bondi-Sachs definition, whereas the concept of angular momentum remains somewhat problematical.

The main ingredient of the Bondi-Sachs energy-momentum definition is a certain curvature quantity, often denoted Ψ_2, which describes a part of the (conformal) curvature whose fall-off at large distances along a null

geodesic is like r^{-3}, where r is an affine parameter on the geodesic. In terms of $\mathcal{I}$, say $\mathcal{I}^+$, we choose a convenient $\hat{g}$ metric for which the metric of $\mathcal{I}^+$ takes the form (4), with u scaled compatibly with the strong conformal geometry, cf. (5), (6), (7). This last may be achieved by taking the physical g-metric in the form

$$\begin{aligned} ds^2 &= 2dudr - r^2(d\theta^2 + \sin^2\theta d\phi^2) \\ &\quad + O(r^{-2})dr^2 + O(r^{-1})\cdot dr(\times \text{ differentials not involving } dr) \\ &\quad + O(r)\,(\times \text{ products of differentials not involving } dr) \end{aligned} \tag{9}$$

where, near $\mathcal{I}^+$, the hypersurfaces $u = \text{const.}$ are taken to be null with r as an affine parameter on each null geodesic generator, and where $\Omega = r^{-1}$ near $\mathcal{I}^+$. Coordinates for which the metric takes the form (9) are called *Bondi-Sachs coordinates* (although usually a slightly different condition, irrelevant for our purposes, is placed on the coordinate r).

It is also convenient, near $\mathcal{I}^+$, to choose a *null tetrad* (with respect to $\hat{g}$) of vectors $\hat{\ell}^a$, $\hat{m}^a$, $\bar{\hat{m}}^a$, $\hat{n}^a$ at points of $\bar{\mathcal{M}}$, and, at points of $\mathcal{M}$, also the null tetrad (with respect to g) defined by

$$\ell^a = \Omega^2\hat{\ell}^a, \quad m^a = \Omega\hat{m}^a, \quad n^a = \hat{n}^a. \tag{10}$$

(The vectors ℓ^a, n^a are real null; and m^a is complex null—in the sense $g_{ab}m^am^b = 0$ —with complex conjugate $\bar{m}^a$. The normalization conditions are $g_{ab}\ell^an^b = 1 = -g_{ab}m^a\bar{m}^b$, the remaining scalar products all vanishing. The conditions on the "hatted" tetrad are the same, but now with respect to $\hat{g}$.) The vector $\hat{n}^a$ is given by (7) on $\mathcal{I}^+$ and

$$g_{ab}\ell^b = \ell_a = \nabla_a u = \hat{\ell}_a = \hat{g}_{ab}\hat{\ell}^b \tag{11}$$

in $\mathcal{M}$ (near $\mathcal{I}^+$). With these conventions, we define the complex tetrad components

$$\begin{aligned} \Psi_0 &= C_{abcd}\ell^am^b\ell^cm^d \\ \Psi_1 &= C_{abcd}\ell^am^b\ell^cn^d \\ \Psi_2 &= C_{abcd}\ell^am^b\bar{m}^cn^d \\ \Psi_3 &= C_{abcd}\ell^an^b\bar{m}^cn^d \\ \Psi_4 &= C_{abcd}\bar{m}^an^b\bar{m}^cn^d \end{aligned}$$

of the Weyl curvature tensor, the corresponding definitions for the "hatted" components also holding, whence, by (10)

$$\Psi_i = \Omega^{i-6}\Psi_i \quad (i = 0,\cdots,4)\,. \tag{12}$$

It can be shown (Penrose 1965b) that for a k-asymptotically simple space-time ($k \geq 3$) which is vacuum near $\mathscr{I}^+$, all the quantities $\hat{\Psi}_i$ have smoothly vanishing limits at $\mathscr{I}^+$ and so

$$\Psi_i = O(r^{i-5}) \quad (i = 0,\cdots,4) \tag{13}$$

which is the "peeling property" of Sachs (1962a, cf. also Newman and Penrose 1962, Penrose 1965b).

We note, from (13), that the quantity Ψ_4 behaves as r^{-1} and so dominates at large distances. It is therefore often referred to as the *gravitational radiation field*. Also the quantity Ψ_2 behaves as r^{-3}, as was alluded to earlier.

We are still not quite in a position to define the Bondi-Sachs mass. For this, we need the further quantity

$$\sigma \quad m^a m^b \nabla_a \ell_b \tag{14}$$

which defines the (complex) *shear* of the null geodesic generators of each null hypersurface $u = \text{const.}$, i.e. of the system of integral curves of ℓ^a. Under conformal rescaling we have

$$\hat{\sigma} = \Omega^{-2}\sigma \tag{15}$$

which implies that

$$\sigma = O(r^{-2})\,,$$

since $\hat{\sigma}$ must be continuous at $\mathscr{I}^+$. A particular cut $u = \text{const.}$ is then defined to be a good cut iff $\hat{\sigma} = 0$ all over it (at $\mathscr{I}$). In Minkowski space, this definition agrees with the one given earlier, but it also allows us to extend the concept of a "good cut" to general asymptotically simple space-times.

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

In fact, for this purpose, the coordinate system that we have introduced is largely an irrelevance. All we need to do is to set up $\hat{\ell}^a$ and $\hat{m}^a$ at each point of the cut (some general cross-section of $\mathscr{I}^+$), where $\hat{m}^a$ is complex null, with real and imaginary parts both tangent to the cut and where $\hat{\ell}^a$ is real null and orthogonal to $\hat{m}^a$ (but not tangent to $\mathscr{I}^+$). The cut is a good cut iff $\hat{\sigma} = 0$, with $\hat{\sigma} = \hat{m}^a\hat{m}^b\hat{\nabla}_a\hat{\ell}_b$ as in (14). In fact σ is the complex shear, at $\mathscr{I}^+$, of that particular null hypersurface in $\mathscr{M}$ which is generated by the null geodesics (other than those in $\mathscr{I}^+$) meeting the cut orthogonally. This is the unique null hypersurface meeting $\mathscr{I}^+$ in the cut in question. In Minkowski space, this null hypersurface is a null cone iff the cut is a good cut.

Another way to think of $\hat{\sigma}$, at points of some general cut $\mathscr{C}$ of $\mathscr{I}^+$, is as a measure of the trace-free part of the extrinsic curvature of $\mathscr{C}$, as it is imbedded in $\mathscr{I}^+$. The modulus of $\hat{\sigma}$ defines the magnitude of this curvature and the argument of $\hat{\sigma}$, the directions of maximum extrinsic curvature (in relation to $\mathrm{Re}(\hat{m}^a)$ and $\mathrm{Im}(\hat{m}^a)$).

Merely having a *definition* of a good cut for general asymptotically flat space-times does not remove the difficulties of defining "supertranslation-free Lorentz rotations" however. For it turns out that when outgoing gravitational radiation is present ($\hat{\Psi}_4 \neq 0$ on $\mathscr{I}^+$), good cuts do not generally exist at all on $\mathscr{I}^+$. The reason for this, roughly speaking, is that $\hat{\sigma}$, being complex, represents two real numbers per point of the cut, whereas to define a particular cut, we have available only one real number per point of the cut (i.e. the "distance" along each generator of $\mathscr{I}^+$). Furthermore, in any Bondi coordinate system the rate of change of $\hat{\sigma}$ (on $\mathscr{I}$), with respect to u, for the family of cuts $u = \text{const.}$ is given by

$$\frac{\partial}{\partial u}\hat{\sigma} = -\bar{N} \tag{16}$$

where N is the Bondi-Sachs (complex) *news function* on $\mathscr{I}^+$, satisfying

$$\frac{\partial}{\partial u}N = \hat{\Psi}_4 \tag{17}$$

showing that the presence of radiation is an obstruction to the Bondi family $u = \text{const.}$ being all good cuts. It may be remarked, also, that N has an additional geometrical interpretation:

$$\bar{N} = -\frac{1}{2}\hat{R}_{ab}\hat{m}^a\hat{\bar{m}}^b$$

(and, of course, $\hat{R}_{ab}$ need not vanish near $\mathscr{I}^+$ even though R_{ab} does).

The news function N is important also in that it features in the definition of the Bondi-Sachs mass at a cut $\mathcal{C}$ of $\mathscr{I}^+$:

$$m = \frac{1}{4\pi}\int(\hat{\sigma}N - \hat{\Psi}_2)\,d\mathcal{C} \tag{18}$$

(which is always real, even though the integrand is complex), the integral being taken over $\mathcal{C}$, where $d\mathcal{C}$ is the element of surface area

$$d\mathcal{C} = \sin\theta\, d\theta \wedge d\phi\ .$$

The original Bondi-Sachs mass-loss formula (Bondi 1960, Bondi et al. 1962, Sachs 1962) can be stated as

$$\frac{dm}{du} = -\frac{1}{4\pi}\int N\bar{N}\,d\mathcal{C} \leq 0$$

where the cuts are taken as the $u = \text{const.}$ surfaces in a Bondi-Sachs coordinate system. This entails that each cut is obtainable from each other cut by a *translation*—in fact by a translation generated by $\partial/\partial u$. But the more general mass-loss formula

$$m_2 - m_1 = -\frac{1}{4\pi}\int_{\mathcal{C}_1}^{\mathcal{C}_2} N\bar{N}\,d\mathcal{C}\wedge du \leq 0 \tag{19}$$

also holds (cf. Penrose 1964, 1967) where m_1 and m_2 simply refer to

two arbitrary cuts $\mathcal{C}_1$ and $\mathcal{C}_2$ of $\mathscr{I}^+$, respectively, where $\mathcal{C}_2$ lies entirely to the future of $\mathcal{C}_1$, the same formula (18) being used in each case. The integral in (19) is taken over the region of $\mathscr{I}^+$ lying between $\mathcal{C}_1$ and $\mathcal{C}_2$, the equation (19) being a consequence of $R_{ab} = 0$ holding in the immediate neighborhood of this region. There is also a generalization of this result when merely the weak energy condition holds in place of Einstein's vacuum equations. An additional term is added to $N\bar{N}$ in the integrand in (19) which is necessarily non-negative whenever (8) holds, so the same inequality applies as before.

It should be clear that the reason we get a mass loss rather than a mass gain is simply that the above discussion has been applied to $\mathscr{I}^+$ rather than $\mathscr{I}^-$. Exactly analogous results to the above would be obtained, as applied now to the Bondi-Sachs mass defined on cuts of $\mathscr{I}^-$, but where now the mass is non-decreasing with time rather than non-increasing. (This assumes that the appropriate regularity and energy conditions are now assumed to hold at $\mathscr{I}^-$ rather than at $\mathscr{I}^+$.) The reason that for physically interesting space-times one is normally concerned with the mass defined at $\mathscr{I}^+$, rather than at $\mathscr{I}^-$, is that physically realistic models would usually involve retarded (i.e. outgoing) radiation but no advanced (i.e. incoming) radiation. The radiation field on $\mathscr{I}^-$ is usually set to zero, in accordance with this, and the Bondi-Sachs mass defined on $\mathscr{I}^-$ is therefore simply a constant. In such circumstances the study of quantities defined at $\mathscr{I}^-$ is less interesting than of the corresponding quantities defined at $\mathscr{I}^+$.

We are now in a position to state several more problems:

PROBLEM 6. Show that if a cut $\mathcal{C}$ of $\mathscr{I}^+$ (or $\mathscr{I}^-$) can be spanned by a spacelike hypersurface along which an appropriate energy condition holds, then the Bondi-Sachs mass defined at $\mathcal{C}$ is non-negative.

The usual energy condition on the energy tensor T_{ab} that is imposed for this sort of problem is the *dominant energy condition* (cf. Synge 1956, Hawking and Ellis 1973):

$$T_{ab}\, t^a u^b \geq 0 \quad \text{whenever } t^a \text{ and } u^a \text{ are future-timelike}$$

which is to be taken in conjunction with Einstein's equation (3). An affirmative resolution of Problem 6 would be a null-infinity version of the Schoen-Yau theorem (1979, 1980). However it is as yet not even completely clear that the two different mass concepts involved here are, in an appropriate sense, the same. If $\mathscr{M}$ admits a spacelike hypersurface which is sufficiently asymptotically flat that a mass can be defined at spacelike infinity according to one of the standard definitions (e.g. Einstein 1918, Arnowitt *et al.* 1962, Landau and Lifshitz 1962) then we ask:

PROBLEM 7. Does the Bondi-Sachs mass defined on cuts of $\mathscr{I}^+$ have a well-defined limit as the cuts recede into the past along $\mathscr{I}^+$, this limit agreeing with the mass defined at spacelike infinity?

The statement of Problem 7 is left deliberately a little vague, it being not quite clear what the appropriate definitions at spatial infinity should be (but cf. Ashtekar and Magnon-Ashtekar 1979). It is necessary, however, that one feature of our mass definitions that I have so far glossed over be now properly taken into account. This is that the mass is not actually a scalar quantity but is one component of an "asymptotic 4-vector," the other three components defining the spatial momentum. The way that the non-scalar nature of the Bondi-Sachs mass, as defined by (18) manifests itself is that there is a certain non-uniqueness in the choice of unit sphere metric that arises in (4), compatible with the given conformal structure. This non-uniqueness shows up in the 3-parameter freedom in the choice of the function $K(\theta, \phi)$ that arises in the BMS transformations. The different choices of $K(\theta, \phi)$ correspond to the different possible choices of "asymptotic time-axis."

If we rescale the unit sphere metric in this way, we obtain a new weighting in the integrand in (18). The 4-dimensional linear space spanned by the factors that so arise provides the asymptotic 4-vector-space for the Bondi-Sachs energy-momentum. (In fact, for each spacelike component, the

resulting weighting factor is negative in places and does not itself arise from a rescaling of the sphere metric.) In a convenient orthonormal basis frame for the asymptotic space, the Bondi-Sachs 4-momentum has components

$$P_a = \frac{1}{4\pi}\int W_a(\hat{\sigma}N - \hat{\Psi}_2)\,d\mathcal{C} \tag{20}$$

with

$$W_0 = 1\,,\quad W_1 = \cos\theta\,,\quad W_2 = \sin\theta\sin\phi\,,\quad W_3 = \sin\theta\cos\phi\,.$$

In fact, an affirmative solution to Problem 6 would imply that P_a is future-causal (i.e. future-timelike, future-null or zero) in the sense

$$P_0 \geq 0\,,\quad (P_0)^2 - (P_1)^2 - (P_2)^2 - (P_3)^2 \geq 0\,. \tag{21}$$

But one would anticipate that, except in the cases when $\mathcal{C}$ spans a spacelike hypersurface entirely lying in flat space-time, the *strict* inequality holds in (21) (assuming satisfaction of the dominant energy condition), i.e. that P_a is future-timelike. Indeed, it follows from Bondi-Sachs mass loss that P_a is future-timelike if it is already so for some later cut of $\mathcal{I}^+$. As far as I am aware, the expected relation (21) at spacelike infinity is not known either, with or without strict inequality.

PROBLEM 8. Show that if the dominant energy condition holds, then the Bondi-Sachs energy-momentum, and also the energy-momentum defined at spacelike infinity, are future-timelike, the space-time being assumed not to be flat everywhere in the region of an appropriate spacelike hypersurface.

The various weighting factors that occur in the definition (20) may be thought of as providing the generators of *translations* of the BMS group. There are other ways of expressing the Bondi-Sachs energy-momentum which makes its relationship to such asymptotic Killing vectors more explicit (Tamburino and Winicour 1966, cf. Winicour 1980). In a corresponding way, one may define a concept of angular momentum at cuts of $\mathcal{I}^+$

whenever a generator of a "Lorentz rotation" belonging to the BMS group has been provided. However, this would require a suitable definition of a "good cut," or something similar, in order to distinguish the six "genuine" components of angular momentum from the infinite-dimensional space of "spurious" components that involve parts arising from generators of supertranslations.

The problem is particularly blatant in the case of (asymptotically simple) space-times for which the outgoing radiation tails off suitably in both directions along $\mathscr{I}^+$, so that "near i^0" and "near i^+" good cuts exist, but where the "good cuts near i^0" and the "good cuts near i^+" are non-trivial supertranslations of one another. Let us say that i^0 *and* i^+ *are non-trivially related* in such space-times. It would seem that the very concept of angular momentum gets "shifted" with time whenever i^0 and i^+ are non-trivially related, so it is hard to see in these circumstances how one can rigorously discuss such questions as the angular momentum carried away by gravitational radiation. Moreover, it would seem that i^0 and i^+ are likely to be non-trivially related in the general case:

PROBLEM 9. In an asymptotically simple space-time which is vacuum near $\mathscr{I}$ and for which outgoing radiation is present and falls off suitably near i^0 and i^+, is it necessarily the case that i^0 and i^+ are non-trivially related? (At least, are there some examples in which i^0 and i^+ are non-trivially related?)

It might be that a somewhat different approach to the problem of angular momentum will evade this problem. Thus we have:

PROBLEM 10. Find a good definition of angular momentum for asymptotically simple space-times.

In this connection, it is worth mentioning an idea due to Newman (1976, cf. Hansen et al. 1978), although a discussion of its many intriguing ramifications would take us too far from our essential purposes. The idea is

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

to allow *complex* cuts of $\mathscr{I}^+$ —or, more accurately, cuts of the complexification $C\mathscr{I}^+$ of $\mathscr{I}^+$. For this purpose one must assume that the space-time $\mathscr{M}$ has an appropriately analytic metric at $\mathscr{I}^+$ (with analytic conformal factor). Then the quantity $\hat{\sigma}$ has a holomorphic extension to these complex cuts, and we can define a complex good cut to be one for which this holomorphically extended $\hat{\sigma}$ vanishes. It can then be shown, assuming that the metric of $\mathscr{M}$ is, in an appropriate sense, "close" enough to that of Minkowski space, that there is a four-complex-dimensional family of such complex good cuts, and, rather remarkably, that the four-complex-dimensional space whose points represent these cuts can, in a natural way, be assigned a complex metric which is automatically Ricci-flat with a self-dual conformal curvature (Hansen et al. 1978, cf. also Penrose 1976). In fact, these complex "space-times" provide *generic* such self-dual solutions of Einstein's vacuum equations—and for this, the vacuum equations need not be assumed to hold for the original metric g_{ab}. Newman's idea leads us into the theory of curve-space *twistors* (Penrose 1976, Penrose and Ward 1980) which provides, among other things, methods of constructing self-dual solutions of Einstein's vacuum equations. A major problem of that theory is to find an appropriate way of extending these techniques which will enable the self-dual restriction to be removed, (cf. Penrose 1979b).

Most of these later problems have been concerned only with outgoing gravitational radiation (or, by time-reflection, only with incoming radiation). There are various problems which are concerned with the interrelations between incoming and outgoing radiation. Let me state, as a mathematical problem, just one of these:

PROBLEM 11. If there is no incoming radiation and no outgoing radiation and the space-time $\mathscr{M}$ is vacuum near $\mathscr{I}$ and (in some suitable sense) near i^0, is $\mathscr{M}$ necessarily stationary near $\mathscr{I}$?

One would anticipate an affirmitive answer to Problem 11. The result of Papapetrou (1957) which asserts the non-existence of asymptotically flat

non-stationary periodic space-times which are Ricci-flat near spatial infinity may be regarded as a step towards a solution to this problem. (A periodic asymptotically simple $\mathscr{I}$ would necessarily be free of both outgoing and incoming radiation, because of (17) and the mass-loss formula (19).) Problem 11 should be considered in conjunction with Problem 4.

There are many other outstanding problems in general relativity involving gravitational radiation, but often they are hard to state simply as mathematical problems. Most notable among these, perhaps, is to obtain an entirely rigorous derivation of Einstein's formula for the radiation emitted by a system of varying quadrupole moment (cf. Landau and Lifshitz 1962). At one level, the study of the structure of $\mathscr{I}^+$ can be said to give a clear-cut derivation (cf. Newman and Unti 1962, Janis and Newman 1965). But the quadrupole moment is defined in terms of quantities that may not be easy to relate directly to local measurable quantities (e.g. star orbits). In the view of some physicists (cf. Ehlers 1980), there could even be some significant doubt about the validity of the Einstein formula. A more substantial problem even than this, however, is to obtain a general understanding of how gravitational waves scatter one another. But this question is rather open-ended. Another problem is to understand the significance (either mathematical or physical) of a certain curious set of ten exactly conserved asymptotically defined quantities (Newman and Penrose 1968, cf. Exton et al. 1969).

I have been discussing problems, for the most part, which refer to asymptotically simple space-times. Roughly, these are space-times in which every light ray escapes (both in past and future directions) to an asymptotically flat region. There are, however, interesting asymptotically flat space-time models in which light rays can get trapped. Most notably, there are the black hole models and models representing dynamical collapse to a black hole, as in Figure 4. In the collapse models there is an initial Cauchy hypersurface $\mathscr{S}$ along which a non-singular space-time geometry is specified. The space-time near $\mathscr{S}$ is taken to be Ricci-flat outside some compact region of $\mathscr{S}$, and inside this region has a Ricci

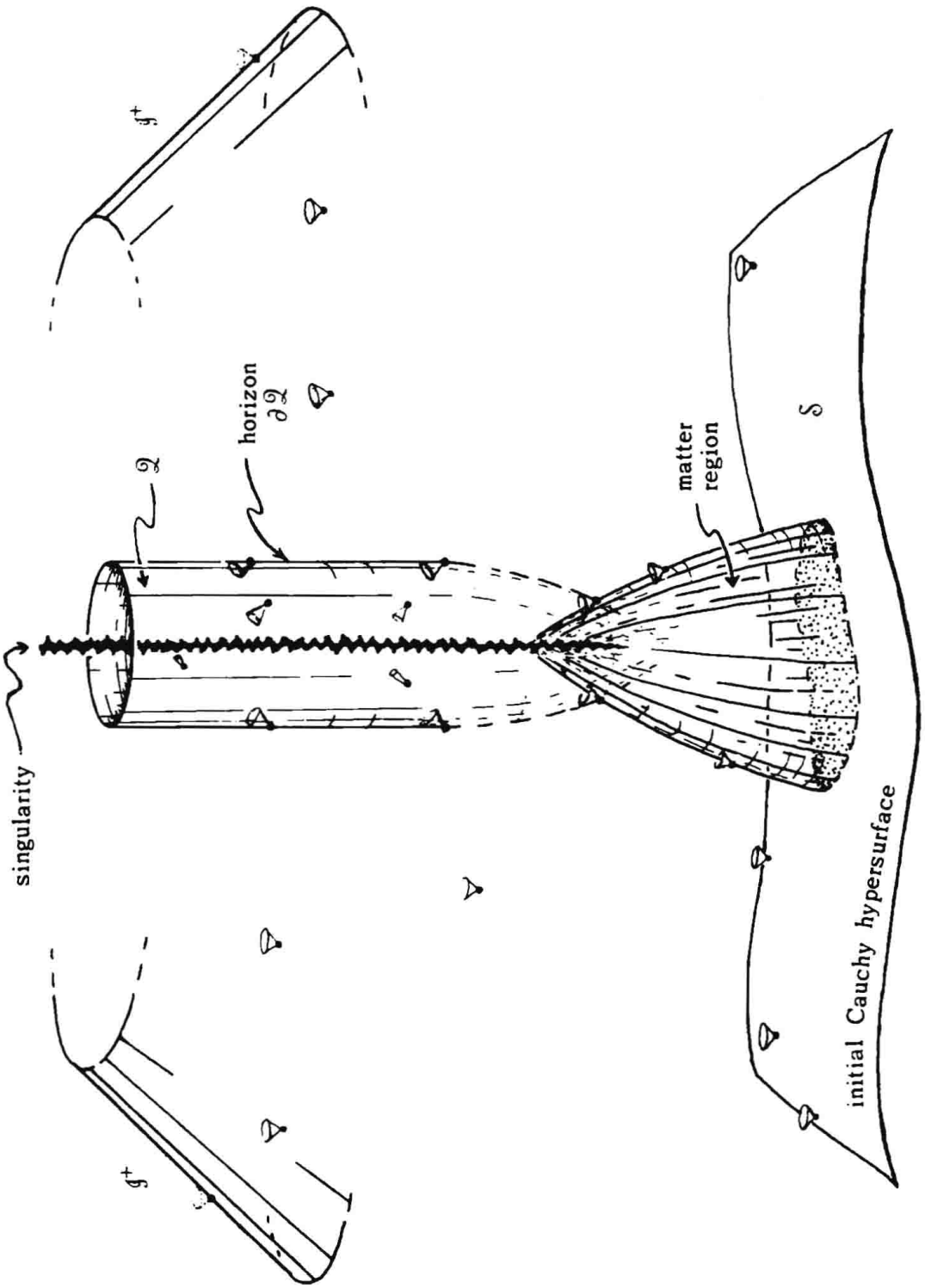


Figure 4. Dynamical collapse to a black hole.

tensor subject to some physically reasonable equation of state (for an appropriate matter distribution). These equations are evolved into the future away from $\mathcal{S}$ and, for certain sets of initial conditions in which the matter region is sufficiently concentrated or falling rapidly enough inwards, one finds that a space-time singularity arises. In this context, by a "space-time singularity arising," we mean that the maximal evolution of the field equations provides a space-time that is not geodesically complete in future-causal (or past-causal) directions—and, indeed, cannot be extended to such a complete space-time even if the field equations are abandoned in the extension. In normal circumstances, the singularity arises because curvature scalars (e.g. $R_{abcd}R^{abcd}$) diverge to infinity at finite affine distance.

It is not hard to produce explicit models of collapse that exhibit this behavior in which spherical symmetry is assumed (cf. Oppenheimer and Snyder 1939). At one time many people believed that the singularity in these models was physically spurious, arising merely because the spherical symmetry implied an exact focussing towards a central point. It was argued that perhaps the introduction of deviations from spherical symmetry would remove this focussing effect, so that infinite curvature would not arise in the central regions. It turned out, however, that this is not the case (Penrose 1965a, Hawking and Penrose 1970, Hawking and Ellis 1973), the formation of such singularities being, in a clear sense, a *stable* property of the initial data sets. Any small-enough (but finite) perturbation of the initial data for spherically symmetrical collapse will also lead to a space-time singularity arising, it being essentially only necessary to assume that the equations of state are such that the *strong energy condition*

$$2\,T_{ab}t^a t^b \geq T^a{}_a t^b t_b \quad \text{whenever} \quad t^a t_a \geq 0$$

always holds, i.e. (with Einstein's equation (3)), we have the *timelike convergence condition*:

$$R_{ab}t^a t^b \leq 0 \quad \text{whenever} \quad t^a t_a \geq 0\,. \tag{22}$$

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

If it assumed that the hypersurface $\mathcal{S}$ remains a Cauchy surface for the entire space-time $\mathcal{M}$ no matter how far $\mathcal{M}$ is extended, then we need only assume null convergence condition (8). A "Cauchy surface" here means a hypersurface with the property that it intersects every maximally extended timelike curve in exactly one point. The hypersurface $\mathcal{S}$, here, will necessarily be a Cauchy surface for the space-time which represents the maximal evolution (cf. Choquet-Bruhat and Geroch 1969) of the initial data on $\mathcal{S}$, but it is conceivable that the space-time could be extended beyond this maximal evolution–to result in a larger space-time for which $\mathcal{S}$ is not a Cauchy hypersurface.

In the spherically symmetrical collapse models there is a certain region $\mathcal{Q}$ composed of those points from which it is not possible to draw a timelike curve of infinite length into the future. Any timelike curve in $\mathcal{Q}$, if extended maximally into the future will encounter the space-time singularity. The boundary $\partial\mathcal{Q}$ of $\mathcal{Q}$ will be referred to here as the *absolute event horizon*, or simply the *horizon*, and the region $\mathcal{Q}$ as a *black hole*.* One anticipates that if the initial data is perturbed slightly away from spherical symmetry, then a picture arises, qualitatively similar to the one just described. In particular, one does not anticipate that timelike curves can escape, in the future directions from the close vicinity of the singularity, to reach the exterior infinity. If the initial data is perturbed by a *large* amount away from spherical symmetry, but in such a way that a singularity still arises, then it is not nearly so clear that the picture will remain essentially unchanged. Perhaps the initial data can evolve into a so-called *naked singularity* from which signals (that is, causal curves) can escape in the future direction to external infinity. This possibility is a disturbing one for astrophysicists, since one has no clear way to make predictions once a naked singularity arises. The normal belief seems to be that naked singularities will not in fact arise–at least if one

*Often a slightly different definition of these concepts is given, where future-timelike curves in $\mathcal{Q}$ do not reach $\mathcal{I}^+$ (see Hawking and Ellis 1973).

assumes that the initial setup is, in some appropriate sense "generic," and that the evolution can be treated as taking place according to standard classical physics. If this is the case, then we say that *Cosmic Censorship* holds (cf. Penrose 1969, 1978, Tipler *et al.* 1980). The proving (or disproving) of an appropriate form of Cosmic Censorship in classical general relativity is perhaps the most important unsolved problem in that theory. It is particularly important because, unlike many of the other problems that I have been discussing, it has genuine astrophysical significance and, furthermore, it is by no means absolutely clear on physical grounds that Cosmic Censorship should always hold true. Thus, we would certainly like to know the answer to:

PROBLEM 12. Is Cosmic Censorship a valid principle in classical general relativity?

The statement of this problem has been left deliberately rather vague. In fact, it is a problem in itself to find a satisfactory mathematical formulation of what is physically intended. For myself, I prefer a rather strong "local" version of Cosmic Censorship which makes little reference to infinity (Penrose 1978, 1979a). The question concerns the structure of the singularities that can arise in the evolution of suitably "generic" initial data on a suitable spacelike hypersurface. Suppose the local structure of these singularities is always such that they provide a *future* boundary to the space-time (for the forward-evolved data, that is; for the backward-evolved data we require the singularities to provide a *past* boundary). Then (in the forward region) there would be no timelike curves going into the future from the vicinity of the singularity to reach external infinity, so we could say that a form of Cosmic Censorship holds. In effect, we can phrase this version of Cosmic Censorship roughly in the following way: Let $\mathcal{S}$ be a complete spacelike initial data surface whose metric approaches that of Euclidean 3-space at infinity and on which are specified initial data (satisfying the necessary constraints) for the gravitational field and for appropriate matter fields. Suppose that these data are evolved into the

 Some unsolved problems in classical general relativity. *Seminar on differential geometry*, 631–68. *Ann. of Math. Stud.*, 102, 1982

future (and past) to obtain the space-time of maximal evolution. The question is then whether this space-time can be extended further, in a way consistent with the field equations. If—at least for "almost all" initial data sets—it is not possible to extend the maximal evolution in this way, then we say that a form of Cosmic Censorship holds. The connection between this kind of extension and the existence of naked singularities is illustrated in Figure 5.

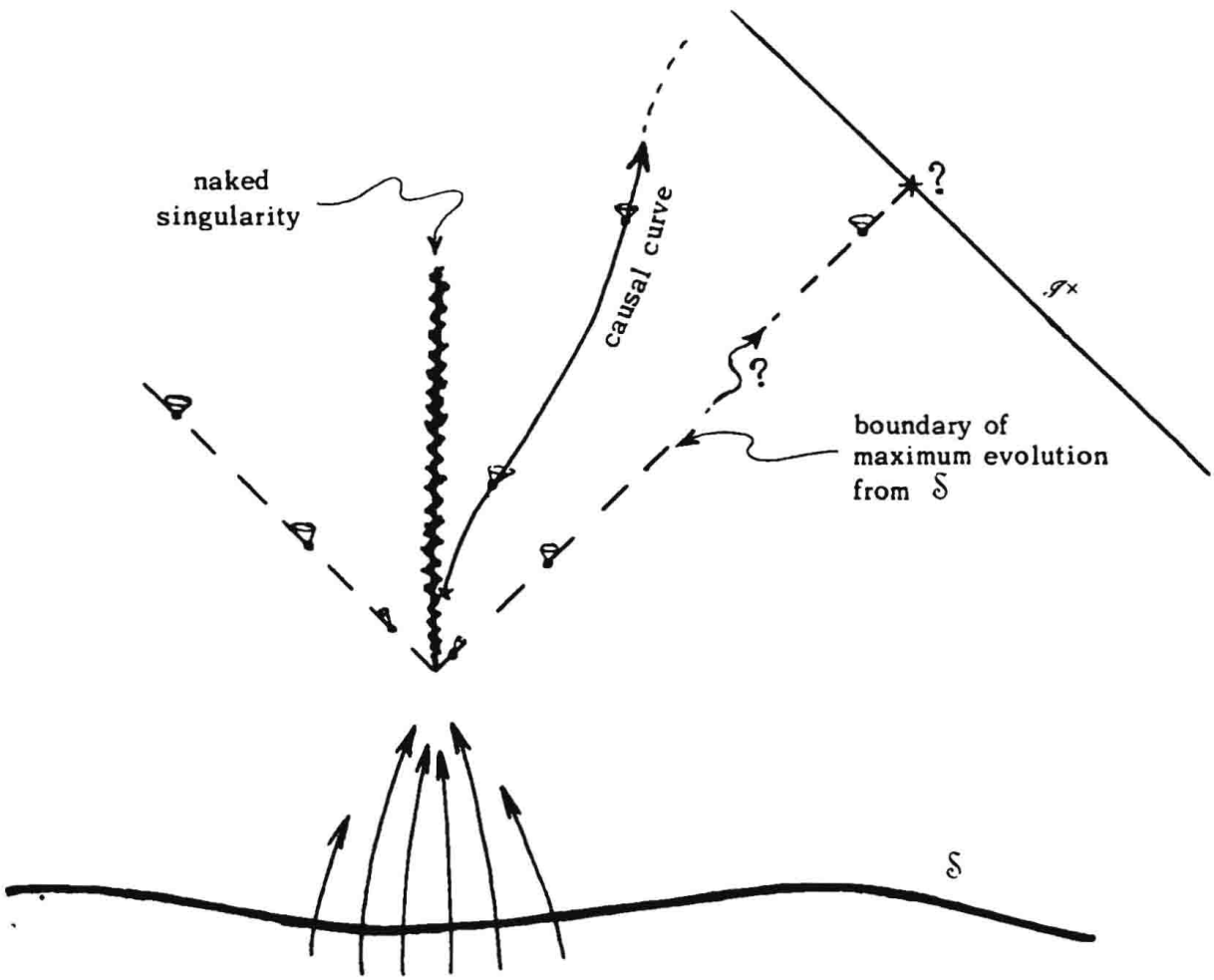


Figure 5. The development of a naked singularity

It may be felt, however, that this version of Cosmic Censorship does not capture all of its physical aspects. One might require, in addition, that provided the initial data tail off appropriately at infinity, the evolution leads to a suitable $\mathscr{I}$ with "infinitely long" generators (cf. Tipler *et al.* 1980). Furthermore, one must be somewhat careful about one's choice of equation of state for the matter region. (For example, it is known that the equations of a pressureless fluid or "dust" will lead to spurious "shell crossing" naked singularities, cf. Yodzis *et al.* 1973.) An important preliminary version of Cosmic Censorship would be the case in which all matter fields are absent, so we are simply exploring the structure of Ricci-flat space-times. The suggestion, then, is that "generic" maximally extended Ricci-flat space-times (which contain asymptotically flat spacelike hypersurfaces) should be *globally hyperbolic* in the sense of Leray (1952). For according to a theorem of Geroch (1967), global hyperbolicity is equivalent to the existence of a Cauchy surface. (See also Penrose 1972.)

There are various curious inequalities that can, in effect, be derived from a suitable assumption of Cosmic Censorship. One of the criteria that can be used for the occurrence of a black hole is the existence of a *trapped surface*, which is a compact spacelike 2-surface $\mathscr{S}$ (normally $\cong S^2$) having the property that both systems of null normals to $\mathscr{S}$ are converging in future directions. (For an ordinary spacelike 2-sphere in Minkowski space, only the ingoing null normals, representing a light flash going inwards from the sphere, are converging. The outgoing null normals represent a light flash coming outwards from the sphere and are diverging.) The essential property of a trapped surface is that the "outgoing" null normals are actually converging. This occurs inside the horizon in the (extended) Schwarzschild space-time and it also occurs inside the horizon of a general black hole (cf. Hawking and Ellis 1973), the singularity in the central region being a consequence of this light-trapping property (and energy conditions, etc.).

Now, with a suitable assumption of Cosmic Censorship, together with various mild other physical assumptions, it is possible to prove a certain

sequence of inequalities (Penrose 1973, cf. Hawking 1972). First, the area A of $\mathcal{S}$ will be less than the area of the horizon $\partial\mathcal{Q}$, at the intersection of $\partial\mathcal{Q}$ with the null hypersurface generated by the geodesics tangent to the ingoing null normals at $\mathcal{S}$. Second, the various cross-sections of $\partial\mathcal{Q}$ (which is a null hypersurface, where non-singular) are non-decreasing in area as they proceed into the future. Third, in accordance with the so-called "no-hair" theorem of Israel-Carter-Hawking-Robinson (cf. Carter 1979) one anticipates (problem!) that the black hole will settle down into a Kerr space-time (cf. Hawking and Ellis 1973), with perhaps other regions containing matter and further black holes (moving away). The area of the horizon for a Kerr hole of mass m and angular momentum am is

$$8\pi m(m + (m^2 - a^2)^{1/2})$$

which is not greater than

$$16\pi m^2$$

(the result for the Schwarzschild metric). The mass measured at infinity at late times, i.e. the Bondi-Sachs mass (18) measured at cuts of $\mathcal{I}^+$ in the future limit $u \to +\infty$ should, on physical grounds, be not less than m (since any other matter should contribute positively). Finally, the mass measured at any earlier cut should, by the mass-loss formula (19), be not less than the future limiting value. Combining these inequalities together (and redefining m), we anticipate an affirmative answer to:

PROBLEM 13. Let $\mathcal{S}$ be a spacelike hypersurface in $\overline{\mathcal{M}}$ which is compact with boundary, the boundary consisting of a cut $\mathcal{C}$ of $\mathcal{I}^+$ together with a trapped surface $\mathcal{T}$. Let m be the Bondi-Sachs mass evaluated at $\mathcal{C}$ and let A be the area of $\mathcal{T}$. Show that

$$A \leq 16\pi m^2$$

provided that the dominant energy condition holds throughout some neighborhood of $\mathcal{S}$.

It should be remarked that when black holes are present, the space-time $\mathfrak{M}$ will not actually be asymptotically simple, since condition (3) of asymptotic simplicity will fail (e.g. for the generators of $\partial\mathcal{Q}$). However this should cause no difficulty, since $\mathcal{I}^+$ can still exist for various weakened versions of asymptotic simplicity (e.g. Penrose 1974, Geroch and Horowitz 1978). All that is required is that conditions (i) and (ii) should hold near $\mathcal{I}^+$, with $\mathcal{C} \cong S^2$. Furthermore, the surface $\mathcal{T}$ need not be assumed to be connected in Problem 13; the effects of the various separate black holes combine together appropriately (cf. Hawking 1972). It may be pointed out, also, that an affirmative answer to Problem 13 will have the consequence that the apparently stronger inequality

$$A \leq 16\pi(m^2 - (p_1)^2 - (p_2)^2 - (p_3)^2)$$

must also hold, where p_a are as in (20), because the inequality in Problem 13 must hold for all allowed scalings for the metric on $\mathcal{I}^+$. Note that Problem 13 generalizes Problem 6. There is also a version of Problem 13 which applies at spacelike infinity (cf. Problem 7) and this has been considered by Schoen and Yau.

Problem 13 may be considered as a sort of consistency check on the various physical assumptions involved. The most doubtful of these is presumably Cosmic Censorship, so inequalities of this nature have been regarded as giving some sort of test of that hypothesis. However, they provide probably a rather weak test, as it seems to me that the Cosmic Censorship hypothesis has been appealed to in only a rather peripheral way. It would be nice to understand more about this, however.

In particular classes of models it is sometimes possible to check the inequality in Problem 13 directly (cf. Penrose 1973, Gibbons 1972, Jang and Wald 1977). Indeed, Gibbons has shown that under certain circumstances, the inequality to be proved follows from Minkowski's inequality for convex surfaces in Euclidean 3-space. Another closely related special case, for a positive function f defined on the (θ,ϕ)-sphere, is

$$4\pi \int f^2 d\mathcal{C} \leq \left\{ \int f(1 - \nabla^2 \log f) d\mathcal{C} \right\}^2 , \tag{23}$$

where $d\mathcal{C} = \sin\theta d\theta \wedge d\phi$ as before, the integrals being taken over the sphere, with Laplacian ∇^2, and where

$$\nabla^2 \log f < 1$$

(cf. Penrose 1973). Gibbons showed that with suitable convexity assumptions, (23) follows from Minkowski's inequality. Perhaps (23) is a well-known result?

There are many other problems involving black holes which have not yet been solved. I shall not dwell on these, but mention one:

PROBLEM 14. Show that there is no vacuum equilibrium configuration involving more than one black hole.

To negate Problem 14, one would require a stationary solution of the vacuum equations that is complete up to its horizon $\partial\mathcal{Q}$, this horizon consisting of two or more disconnected pieces. Work by Gibbons, D. C. Robinson, Lindenblom and others has established Problem 14 affirmatively in special cases, but the general result remains unknown. If instead of the vacuum equations we allow a matter density given by a Maxwell field (Einstein-Maxwell equations) then we obtain an analogue of Problem 14 which is not true if we allow certain "degenerate" charged black holes for which electrostatic repulsion just balances the gravitational attraction (Israel and Wilson 1972, Hartle and Hawking 1972). It is not known whether there is a purely gravitational analogue of this, where spin takes over the role of electric charge. This seems rather unlikely. But generally results for the Einstein-Maxwell equations follow closely those for the pure Einstein vacuum equations. Thus Einstein-Maxwell analogues exist for the various other problems that I have stated only in the Einstein vacuum case, namely Problems 3, 4, 9 and 11.

I am grateful to the Institute for Advanced Study, Princeton, for their hospitality and to the National Science Foundation for support under Contract MCS 79-12938.

REFERENCES

Arnowitt, R., Deser, S. and Misner, C.W., (1962), in *Gravitation, and Introduction to Current Research* (ed. L. Witten, Wiley, New York).

Ashtekar, A. and Magnon-Ashtekar, A., (1979), *Phys. Rev. Lett.* 43, 181.

Bondi, H. (1960), *Nature*, Lond. 186, 535.

Bondi, H., van der Burg, M.G.J. and Metzner, A.W.K. (1962), Proc. Roy. Soc. (Lond.) A269, 21.

Carter, B. (1979), in *General Relativity, an Einstein Centenary Survey* (eds. S.W. Hawking and W. Israel, Camb. Univ. Press).

Choquet-Bruhat, Y. and Geroch, R. (1969), Commun. Math. Phys. 14, 329.

Ehlers, J. (1980), *Ann. N.Y. Acad. Sci.*

Einstein, A. (1918), *S. B. preuss Akad. Wiss.* 448.

Exton, A.R., Newman, E.T. and Penrose, R. (1969), *J. Math. Phys.* 10, 1566.

Geroch, R. (1967), *J. Math. Phys.* 8, 782.

———, (1971), in *General Relativity and Cosmology*, International School of Phys. "Enrico Fermi" 47 (Ed. R.K. Sachs, Acad. Press).

Geroch, R. and Horowitz, G. (1978), *Phys. Rev. Lett.* 40, 203; 40, 483.

Gibbons, G.W. (1972), *Commun. Math. Phys.* 27, 87.

———, (1974), *Commun. Math. Phys.* 35, 13.

Hartle and Hawking (1972), *Commun. Math. Phys.* 26, 87.

Hansen, R.O., Newman, E.T., Penrose, R. and Tod, K.P. (1978), *Proc. Roy. Soc.* (Lond.) A363, 445.

Hawking, S.W. (1972), *Commun. Math. Phys.* 25, 152.

Hawking, S.W. and Ellis, G.F.R. (1973), *The Large Scale Structure of Space-Time* (Camb. Univ. Press).

Hawking, S.W. and Penrose, R. (1970), *Proc. Roy. Soc. Lond.* A314, 529.

Israel, W. and Wilson, G.A. (1972), *J. Math. Phys.* 13, 865.

Jang, P.S. and Wald, R. (1977), *J. Math. Phys.* 18. *41.*

Janis, A.I. and Newman, E. (1965), *J. Math. Phys.* 6, 902.

Landau and Lifshitz (1962), *The Classical Theory of Fields* (Trans. M. Hamermesh, Pergamon, Oxford).

Leray, J. (1952), *Hyperbolic Differential Equations* (Princeton, I.A.S.).

Newman, E. T. (1976), *Gen. Rel. Grav.* 7, 107.

Newman, E. T. and Unti, T.W.J. (1962), *J. Math. Phys.* 3, 891.

Newman, E. T. and Penrose, R. (1962), *J. Math. Phys.* 3, 566; 4, 998.

———, (1968), *Proc. Roy. Soc.* A305, 175.

Oppenheimer, J. R. and Snyder, H. (1939), *Phys. Rev.* 56, 455.

Papapetrou, A. (1957), *Ann. Phys.* 20, 399.

Penrose, R. (1964), in *Relativity Groups and Topology: the 1963 Les Houches lectures* (eds. C. M. deWitt and B. S. deWitt, Gordon and Breach, New York).

———, (1965a), *Phys. Rev. Lett.* 14, 57.

———, (1965b), *Proc. Roy. Soc.* A284, 159.

———, (1967), in *Relativity Theory and Astrophysics I* (ed. J. Ehlers, Amer. Math. Soc.).

———, (1969), *Rivista del Nuovo Cim. Numero Speciale* 1 (ser. 1), 252.

———, (1972), *Techniques of Differential Topology in Relativity* (S.I.A.M., Philadelphia).

———, (1973), *Ann. N.Y. Acad.* 224, 125.

———, (1974), in *Group theory in non-linear problems* (ed. A. O. Barut, Reidel, Dordrecht).

———, (1976), *Gen. Rel. Grav.* 7, 31.

———, (1978), in *Theoretical Principles in Astrophysics and Relativity* (eds. N. R. Lebovitz, W. H. Reid and P. O. Vandervoort, Univ. of Chicago Press).

———, (1979a), in *General Relativity, and Einstein Centenary Survey* (eds. S. W. Hawking and W. Israel, Camb. Univ. Press).

———, (1979b), in *Advances in Twistor Theory* (eds. L. P. Hughston and R. S. Ward), Pitman, San Francisco.

Penrose, R. and Ward, R. S. (1980), in *General Relativity and Gravitation*, Vol. 2 (ed. A. Held, Plenum, New York).

Porrill, J. and Stewart, J. M. (1980), *Proc. Roy. Soc.* (Lond.) A

Sachs, R. K. (1962a), *Proc. Roy. Soc.* (Lond.) A270, 103.

———, (1962b), *Phys. Rev.* 128, 2851.

Schmidt, B. G. and Stewart, J. M. (1979), *Proc. Roy. Soc.* (Lond.) A367, 503.

Schoen, R. and Yau, S.-T. (1979), *Commun. Math. Phys.* 65, 45.

———, (1981), *Commun. Math. Phys.* 79, 231-260.

Synge, J. L. (1956), *Relativity, the Special Theory* (North-Holland, Amsterdam).

Tamburino, L. and Winicour, J. (1966), *Phys. Rev.* 150, 1039.

Tipler, F. J., Clarke, C.J.S. and Ellis, G.F.R. (1980), in *General Relativity and Gravitation Vol. 2* (ed. A. Held, Plenum, New York).

Winicour, J. (1980), in *General Relativity and Gravitation, Vol. 2* (ed. A. Held, Plenum, New York).

Yodzis, P., Seifert, H.-J. and Muller zum Hagen, H. (1973), *Commun. Math. Phys.* 34, 135.

Added in proof:

A new line of argument, due to E. Witten, has provided a greatly simplified proof of the Schoen-Yau theorem. It seems that Problems 6 and 8 are also likely to succumb to this line of argument.

132
QUASI-LOCAL MASS AND ANGULAR MOMENTUM IN GENERAL RELATIVITY

Proc. Roy. Soc. London Ser. A **381** (1982), no. 1780, 53–63

This paper (together with related work) appears to have stimulated quite a lot of interest in the possibility of a 'quasi-local' definition of *mass-energy* and/or *angular momentum* in general relativity. For a long time it had seemed that these fundamental physical notions could apply in an exact sense either only at a single location (where the energy-momentum tensor would supply what is needed) or globally (where an asymptotic expression such as the ADM or Bondi-Sachs could apply). A quasi-local definition should apply to a closed (usually spacelike) 2-surface $\mathcal{S}$, purporting to measure the total surrounded by $\mathcal{S}$, where the definition should depend only on the fields and space-time geometry in the immediate neighbourhood of $\mathcal{S}$. The specific definition proposed here passed many tests (see, for example, Chapters 146, 152, 153), but then began to show its limitations (see Chapters 150, 161, 173). Various modifications have been proposed, but as things stand, the status of this idea remains tantalizing and not fully clear.

202 Quasi-local mass and angular momentum in general relativity. *Proc. Roy. Soc. London Ser. A* **381** (1982), no. 1780, 53–63

Quasi-local mass and angular momentum in general relativity

By R. Penrose, F.R.S.

Mathematical Institute, 24–29 St Giles, Oxford OX1 3LB, U.K.

(*Received* 30 *November* 1981)

A new approach to defining energy-momentum and angular momentum in general relativity is presented which avoids some of the difficulties of previous definitions and which can be applied quasi-locally. It depends on the construction of a twistor space $\mathbb{T}^\alpha(\mathscr{S})$ associated with any spacelike topological 2-sphere $\mathscr{S}$. Though several problems of interpretation remain to be solved, the new definition works well at $\mathscr{I}^+$, reproducing the Bondi-mass-momentum as four of the ten precisely determined quantities at each cut of $\mathscr{I}^+$. The remaining six quantities provide a definition of angular momentum which appears to be new.

Introduction

It is perhaps ironic that *energy conservation*, a paradigmatic physical concept arising initially from Galileo's (1638) studies of the motion of bodies under gravity, and which now has found expression in the (covariant) equation

$$\nabla_a T^{ab} = 0 \tag{1}$$

– a cornerstone of Einstein's (1915) general relativity – should nevertheless have found no universally applicable formulation, within Einstein's theory, incorporating the energy of gravity itself. The energy tensor T^{ab}, providing the right-hand side to the Einstein field equation, describes the complete local energy, this being the sum of the energy densities of all *non*-gravitational fields. Gravitational (field) energy, on the other hand, contributes non-locally to the total energy, its presence being manifested in the fact that (1) does not, by itself, give rise to an integral conservation law. To do so, (1) would have had to have had the form of a divergence of a vector – like the equation expressing conservation of electric charge:

$$\nabla_a J^a = 0 \tag{2}$$

– rather than of a valence-2 tensor. The non-local gravitational field energy has, moreover, found clear-cut formulation *only* in the expressions used at infinity for an *asymptotically flat* space–time $\mathscr{M}$ like the A.D.M. mass (Arnowitt *et al.* 1962) at spacelike infinity and the Bondi–Sachs mass (Bondi 1960; Bondi *et al.* 1962; Sachs 1962) at null infinity $\mathscr{I}^+$, and only for the Bondi–Sachs mass is there a precise flux law relating mass loss to the non-local energy stored in gravitational radiation

(Bondi 1960; Bondi *et al.* 1962; Sachs 1962; Penrose 1967*a*, 1968*a*; Newman & Penrose 1968).

Part of the difficulty stems from the fact that mass-energy, being but one component of an energy-momentum 4-vector, does not sum together at different points in a natural way, there being no natural global parallelism of 4-vectors in general space–times. At infinity this problem can be circumvented, there being a well defined (global) concept of an asymptotic 4-vector.

The problem for angular momentum is more serious than this. In flat space–time $\mathbb{M}$, we have a skew tensor M^{ab} with an *origin-dependence* governed by the 4-momentum p_a. The entire structure turns out to be appropriately described by the symmetric *kinematic twistor* $A_{\alpha\beta}$ (Penrose & MacCallum 1973; Penrose & Ward 1980). This origin-dependence causes difficulties even at infinity, for a curved $\mathscr{M}$, and, hitherto, no satisfactory way of circumventing these problems has emerged.

In the present note, a new approach to these questions is outlined. A definition of angular momentum at $\mathscr{I}^+$ is provided which does not suffer from some of the drawbacks of those previously existing (cf. Tamburino & Winicour 1966; Bramson 1975*b*, 1978; Lind *et al.* 1972; Prior 1977; Streubel 1978; Winicour 1968, 1980; Geroch & Winicour 1981), with their inappropriate involvement with 'supermomenta' and (except for the Bramson–Streubel definition) anomalous factors of 2. When specialized to the 4-momentum components, the present prescription concurs precisely with that of Bondi–Sachs (Bondi 1960; Bondi *et al.* 1962; Sachs 1962; cf. also Bonnor & Rotenberg 1966). Moreover, it can be applied also to a *finite* spacelike 2-surface $\mathscr{S}$ to provide a *quasi-local* definition of the ten components of momentum–angular-momentum surrounded by $\mathscr{S}$, though as yet these components emerge as complex, with no manifest reduction to ten real components in the general case. For the quasi-local rest-mass, a clear-cut real formula can be given, though there remains a question as to whether certain alternative definitions might be more meaningful.

Basic procedure

To understand the basis of the present approach, consider the relation between equations (1) and (2). If we set (with T^{ab} symmetric)

$$J^a = T^{ab} k_b, \tag{3}$$

where k_b satisfies the Killing equation

$$\nabla_{(a} k_{b)} = 0 \tag{4}$$

(and where J^a does not now describe electric charge) then we find that (2) is a consequence of (1). In $\mathbb{M}$, there are ten independent solutions of (4) (one for each Poincaré generator), and the various choices of J^a so obtained provide ten independent conservation laws: namely special-relativistic energy-momentum and angular momentum. These are conservation laws which have the form of a flux theorem: the net flux of 'charge' crossing the (closed) 3-boundary of a compact

4-volume is zero. But we may also refer to the total 'charge' η surrounded by (or 'threading through') a closed 2-surface $\mathscr{S}$, namely

$$\eta = \oint_{\mathscr{S}} H_{ab}\, \mathrm{d}x^a \wedge \mathrm{d}x^b, \tag{5}$$

where

$$\nabla_{[a} H_{bc]} = \tfrac{4}{3}\pi\, e_{abcd} J^d. \tag{6}$$

(In the case of electromagnetism, with field tensor F_{ab}, (6) arises in the form $\nabla_a F^{ab} = 4\pi J^b$, where $H_{ab} = F^*_{ab} = \frac{1}{2} e_{abcd} F^{cd}$. Here $e_{abcd} = e_{[abcd]}$, with $e_{0123} = 1$ in a right-handed orthonormal frame. The metric signature is $+ - - -$.)

In weak-field gravitational theory there is an analogue of (5) (compare also Sachs & Bergmann 1958). The weak-field curvature tensor K_{abcd} is a tensor field in $\mathbb{M}$ satisfying

$$K_{abcd} = K_{[cd][ab]}, \quad K_{[abc]d} = 0; \quad \nabla_{[a} K_{bc]de} = 0 \tag{7}$$

and related to T_{ab} by

$$K_{acb}{}^c - \tfrac{1}{2} g_{ab} K_{cd}{}^{cd} = -8\pi G T_{ab}. \tag{8}$$

We seek a skew tensor S^{ab} with the property that, for a given Killing vector k_a,

$$\nabla_{[a}(K_{bc]qd} S^{qd}) = \tfrac{4}{3}\pi G\, e_{abcd} T^{df} k_f. \tag{9}$$

For then, comparison with (6) shows that if we set

$$H_{ab} = \frac{1}{G} K_{abcd} S^{cd}, \tag{10}$$

and substitute into (5) we obtain an integral formula which provides, from the components of K_{abcd} at $\mathscr{S}$, an expression for the total energy, momentum or angular momentum surrounded by $\mathscr{S}$, the different possibilities arising from the different possible choices of S^{ab}. The conditions on $S^{ab} = S^{[ab]}$ (via some non-obvious manipulations) are

$$\nabla^{(a} S^{b)c} - \nabla^{(a} S^{c)b} + g^{a[b} \nabla_d S^{c]d} = 0 \tag{11}$$

and

$$e_{abcd} \nabla^b S^{cd} = \tfrac{3}{2} k_a, \tag{12}$$

and it turns out that there are solutions of (11), (12) in $\mathbb{M}$, for each choice of Killing vector k^a. Moreover (11) by itself has 20 independent solutions, each of which automatically defines a Killing vector in (12), ten independent solutions yielding zero in (12). (Equation (11) is perhaps most easily established by showing first that it is equivalent to the spinor equation (13); also the derivation of (12) is facilitated by the observation that T_{ab} is proportional to a contraction of the double dual of K_{abcd}.)

The system of equations (11), being highly over-determined, has no non-trivial solution in a general curved space–time $\mathscr{M}$. This relates to the fact that (4), also, will have only the trivial solution for general $\mathscr{M}$. While it is thus too much to ask that these equations have solutions throughout $\mathscr{M}$, it is more plausible that, in some sense, the 'restriction of the equations to $\mathscr{S}$' might be soluble. But simply taking parts of the equations tangential to $\mathscr{S}$ does not work, and a more subtle

procedure is required. This involves building S^{ab} and k_a out of a more basic spinor-type object defined at $\mathscr{S}$, referred to here as a *2-surface twistor*.

We adopt standard 2-spinor (abstract-index) conventions, as in Penrose (1968*a*). Then (11) translates as

$$\nabla^{(A}_{A'}\gamma^{BC)} = 0, \tag{13}$$

where

$$S^{ab} = \gamma^{AB}\epsilon^{A'B'} + \epsilon^{AB}\bar{\gamma}^{A'B'} \tag{14}$$

with

$$\gamma^{AB} = \gamma^{BA}. \tag{15}$$

Equation (13) is the valence-$\begin{bmatrix}2\\0\end{bmatrix}$ symmetric *twistor equation* (cf. Penrose 1967*b*) and it has ten independent complex solutions in $\mathbb{M}$. Its main significance here lies in the fact that for a symmetric spinor ϕ_{ABCD}

$$\nabla^{A}_{A'}\phi_{ABCD} = 0 \quad \textit{implies} \quad \nabla^{AA'}(\phi_{ABCD}\gamma^{CD}) = 0, \tag{16}$$

(cf. Penrose 1968*b*). For weak-field gravity, where $T_{ab} = 0$, K_{abcd} takes the form

$$K_{abcd} = \phi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\bar{\phi}_{A'B'C'D'}, \tag{17}$$

where ϕ_{ABCD} is symmetric and satisfies the massless spin-2 equation $\nabla^{A}_{A'}\phi_{ABCD} = 0$. Combining (14) with (17) and substituting in (10) we find

$$\tfrac{1}{2}GH_{ab} = \phi_{AB}\epsilon_{A'B'} + \epsilon_{AB}\bar{\phi}_{A'B'}, \tag{18}$$

where

$$\phi_{AB} = \phi_{ABCD}\gamma^{CD} \tag{19}$$

satisfies, by (16), the massless spin-1 equation $\nabla^{A}_{A'}\phi_{AB} = 0$. This tells us that H_{ab} satisfies the free-field Maxwell equations, so the right-hand side of (6) vanishes, as required for (9), where $T_{ab} = 0$.

This (spinor) method of reducing the spin-2 conservation laws (linear gravity) to those for spin-1 (Maxwell field) seems to have been proposed first by the author (Penrose 1968*b*) and was followed up by Streubel (1978) in an attempt to define angular momentum in full general relativity at null infinity $\mathscr{I}^+$. However, Streubel's method of defining 'γ^{AB}' on a general cut (cross section of $\mathscr{I}^+$) is not quite appropriate; it allows too much latitude, so that the 'supermomenta' are not eliminated. The essential problem with using (13) (or equivalently (11)) directly on $\mathscr{S}$ is that, of the eight complex equations in (13), only two are entirely tangential to $\mathscr{S}$, which is insufficient for determining the three independent (complex) components of γ^{AB}.

To circumvent this problem we think of the defining equation (13) for γ^{AB} in a different light. Any solution of (13) in $\mathbb{M}$ has the form

$$\gamma^{AB} = \sum_{i} \omega^{(A}_{i}\omega^{B)}_{j}, \tag{20}$$

where each ω^{A}_{i} is a solution of the (valence-$\begin{bmatrix}1\\0\end{bmatrix}$) *twistor equation*

$$\nabla^{(A}_{A'}\omega^{B)} = 0. \tag{21}$$

(In fact the sum in (20) need contain only two terms, so the general γ^{AB} can be constructed from just four solutions of (21).) Taking $\mathscr{S}$ as a general spacelike 2-surface (with S^2 topology) in a curved $\mathscr{M}$, we can choose spin-frames (o^A, ι^A) at the

points of $\mathscr{S}$ so that the null vectors $l^a = o^A o^{A'}$ and $n^a = \iota^A \iota^{A'}$ are orthogonal to $\mathscr{S}$. This is an appropriate set-up for the G.H.P. formalism (Geroch *et al.* 1973), which eliminates the necessity of making a specific choice of frame. We find that there are two entirely tangential equations among the six G.H.P. equations defined by (21), these being

$$\eth' \omega^0 = \sigma' \omega^1, \quad \eth \omega^1 = \sigma \omega^0. \tag{22}$$

It can be shown, by using the Atiyah–Singer index theorem, that the coupled equations (22) always have at least four linearly independent solutions (ω^0, ω^1), and that for 'normal' situations (for example those that are sufficiently close to the 'canonical case' in which $\mathscr{S}$ is the intersection of two light cones in $\mathbb{M}$, their vertices being timelike-separated), there are *precisely* four independent solutions. (The argument is to show that in this canonical case, where $\sigma = \sigma' = 0$, the adjoint equations to (22) have *no* solution, as follows because the spin-weights are reversed, while equations (22) have *four* solutions. From the invariance of the Atiyah–Singer index, (22) can, at worst, acquire *additional* solutions when σ and σ' become non-zero.) Moreover, equations (22) are conformally invariant (with $o^A \mapsto \Omega^{-\frac{1}{2}} o^A$, $\iota^A \mapsto \Omega^{-\frac{1}{2}} \iota^A$ and $\omega^A \mapsto \omega^A$ under $g_{ab} \mapsto \Omega^2 g_{ab}$, whence $\omega^1 \mapsto \Omega^{\frac{1}{2}} \omega^1$, $\omega^0 \mapsto \Omega^{\frac{1}{2}} \omega^0$).

The space of solutions of (22) will be called the 2-*surface twistor space* with respect to $\mathscr{S}$ and denoted $\mathbb{T}^\alpha(\mathscr{S})$. Thus $\mathbb{T}^\alpha(\mathscr{S})$ is a complex vector space which, in the 'normal' situation, will be four-dimensional. The symmetric tensor product $\mathbb{T}^{(\alpha\beta)}(\mathscr{S})$ of $\mathbb{T}^\alpha(\mathscr{S})$ with itself will then be complex ten-dimensional. Each element of $\mathbb{T}^\alpha(\mathscr{S})$ is a spinor field ω^A defined over $\mathscr{S}$, and taking the symmetric tensor product pointwise on $\mathscr{S}$, we find that each element of $\mathbb{T}^{(\alpha\beta)}(\mathscr{S})$ is a symmetric spinor field γ^{AB} defined over $\mathscr{S}$. Each of the ten linearly independent complex γ^{AB}-fields so obtained may then be substituted into (14) and thence into (10), with the full curvature R_{abcd} now replacing K_{abcd}, and finally into (5) to obtain the various 'charges' surrounded by $\mathscr{S}$. In the limiting case of weak gravitational fields, it follows from the earlier discussion that we obtain correct expressions for the energy-momentum and angular momentum of the sources.

The kinematic twistor

In terms of the standard spin-coefficient notation for the curvature (Newman & Penrose 1962) we can rewrite (5) as a sum of terms, using (20), which are the real parts of expressions

$$\zeta = -\frac{\mathrm{i}}{4\pi G} \oint \{(\Psi_1 - \Phi_{01}) \underset{1}{\omega^0} \underset{2}{\omega^0} + (\Psi_2 - \Phi_{11} - \Lambda) (\underset{1}{\omega^0} \underset{2}{\omega^1} + \underset{1}{\omega^1} \underset{2}{\omega^0}) + (\Psi_3 - \Phi_{21}) \underset{1}{\omega^1} \underset{2}{\omega^1}\} \mathscr{S}, \tag{23}$$

where $\mathscr{S}$ is the surface-area 2-form on $\mathscr{S}$, and $(\underset{1}{\omega^0}, \underset{1}{\omega^1})$ and $(\underset{2}{\omega^0}, \underset{2}{\omega^1})$ are two solutions of (22). There are ten complex linearly independent expressions (23) and we might anticipate that some Hermiticity property of the curvature enables them to be reduced to ten real ones. Indeed, for the analogous situation with electromagnetism, a single complex integral arises whose real and imaginary parts are the electric and

magnetic charge, respectively, whence the *vanishing* of magnetic charge entails that here the complex integral reduces to a real one. With linear gravity there is a corresponding, but more complicated, Hermiticity property which arises as follows.

Rewrite (23) as

$$\zeta = A_{\alpha\beta} \underset{1}{Z^{\alpha}} \underset{2}{Z^{\beta}}, \tag{24}$$

where $\underset{1}{Z^{\alpha}}$ and $\underset{2}{Z^{\beta}}$ are the elements of $\mathbb{T}^{\alpha}(\mathscr{S})$ corresponding to $\underset{1}{\omega^{A}}$ and $\underset{2}{\omega^{A}}$, respectively ($\alpha, \beta, \ldots$, being four-dimensional 'twistor' indices labelling the four-dimensional solution space of (22)). In linear gravity the space–time is $\mathbb{M}$ and $\mathbb{T}^{\alpha}(\mathscr{S})$ can be identified with its standard twistor space $\mathbb{T}^{\alpha}$ of solutions of (21) (Penrose 1967*b*, Penrose & MacCallum 1973, Penrose & Ward 1980), $A_{\alpha\beta} \in \mathbb{T}_{(\alpha\beta)}$ being the total *kinematic twistor* of the sources within $\mathscr{S}$, comprising the four energy-momentum and six angular momentum components. The space $\mathbb{T}_{(\alpha\beta)}$, being the symmetric tensor product of dual twistor space $\mathbb{T}_{\alpha}$ with itself, is ten-complex-dimensional. However, $A_{\alpha\beta}$ satisfies the Hermiticity condition

$$\overline{A_{\alpha\gamma} I^{\beta\gamma}} = A_{\beta\gamma} I^{\alpha\gamma}, \tag{25}$$

where the bar on the left denotes standard twistor complex conjugation, $I^{\alpha\beta}$ being the infinity twistor. The condition (25) reduces the freedom in $A_{\alpha\beta}$ from that of ten complex to ten real components.

For a corresponding reduction to hold in full general relativity we seem to require an analogue of $I^{\alpha\beta}$ in $\mathbb{T}^{\alpha\beta}(\mathscr{S})$ (which is presumably skew, i.e. in $\mathbb{T}^{[\alpha\beta]}(\mathscr{S})$) and also an involutory complex conjugation operation taking $\mathbb{T}^{\alpha}(\mathscr{S})$ to its dual:

$$\mathbb{T}^{\alpha}(\mathscr{S}) \to \mathbb{T}_{\alpha}(\mathscr{S}). \tag{26}$$

Equivalent to (26) is the existence of a Hermitian (bilinear) norm on $\mathbb{T}^{\alpha}(\mathscr{S})$ (signature $+ + - -$) analogous to the standard twistor space norm $Z^{\alpha}\bar{Z}_{\alpha}$. One suggested definition for this norm is

$$Z^{\alpha}\bar{Z}_{\alpha} = \left(\oint \mathscr{S}\right)^{-1} \oint 4 \operatorname{Im}(\omega^{0}\, \eth \bar{\omega}^{1'})\, \mathscr{S} \tag{27}$$

and another is

$$Z^{\alpha}\bar{Z}_{\alpha} = \frac{1}{2\pi} \oint k \operatorname{Im}(\omega^{0}\, \eth \bar{\omega}^{1'} - \bar{\omega}^{1'}\, \eth \omega^{0})\, \mathscr{S}, \tag{28}$$

where k could be the Gaussian curvature

$$k = 2 \operatorname{Re}(\sigma\sigma' - \rho\rho' - \Psi_{2} + \Phi_{11} + \Lambda), \tag{29}$$

of $\mathscr{S}$ (with notation as in Geroch *et al.* 1973), or possibly instead

$$k = \sigma\sigma' + \overline{\sigma\sigma}' - 2\rho\rho', \tag{30}$$

for example. All three of these definitions agree with the standard one for $\mathbb{T}^{\alpha}$ when $\mathscr{S}$ is any allowable surface (spacelike topological 2-sphere) in $\mathbb{M}$. Also the first two agree with the standard definition if $\mathscr{S}$ lies in any conformally flat $\mathscr{M}$. These properties can be derived from the fact that in such cases the expression

$$Z^{\alpha}\bar{Z}_{\alpha} = \omega^{A}\bar{\pi}_{A} + \pi_{A'}\bar{\omega}^{A'}, \tag{31}$$

is *constant* over $\mathscr{S}$, where

$$\pi_{0'} = \mathrm{i}\,\eth'\omega^1 - \mathrm{i}\rho\omega^0, \quad \pi_{1'} = \mathrm{i}\,\eth\omega^0 - \mathrm{i}\rho'\omega^1, \tag{32}$$

the equations (32) being, with (22), restrictions to $\mathscr{S}$ of the standard twistor relation, in $\mathbb{M}$:

$$\nabla_{AA'}\omega^B = \mathrm{i}\epsilon_A{}^B\pi_{A'}. \tag{33}$$

The integrand in (27) is effectively (31), and that in (28) differs from (31) by the factor $\frac{1}{2}k$, which integrates to 4π by the Gauss–Bonnet theorem.

In passing, it may be remarked (as was first pointed out to me by K. P. Tod) that (23) can be re-expressed as

$$\zeta = -\frac{\mathrm{i}}{4\pi G}\oint (\pi_{\underset{1}{0'}}\pi_{\underset{2}{1'}} + \pi_{\underset{1}{1'}}\pi_{\underset{2}{0'}})\,\mathscr{S}, \tag{34}$$

where (32), (22) and equations (2.21), (3.12), (3.13) from Geroch *et al.* (1973) have been used. The full significance of the simple-looking expression (34) remains obscure, however, and no satisfactory analogue of (25) has emerged as yet (though tantalizingly, the flat-space expression for $I_{\alpha\beta}\underset{1}{Z}{}^{\alpha}\underset{2}{Z}{}^{\beta}$ is $\pi_{\underset{1}{0'}}\pi_{\underset{2}{1'}} - \pi_{\underset{1}{1'}}\pi_{\underset{2}{0'}}$).

A definition of the *rest-mass* $m(\mathscr{S})$ surrounded by $\mathscr{S}$ (the quasi-local mass) can be given *without* a full analogue of (25), only the norm defining $Z^{\alpha}\bar{Z}_{\alpha}$ being required. For this, we take over the relation

$$m^2 = -\tfrac{1}{2}A_{\alpha\beta}\bar{A}^{\alpha\beta},$$

from standard twistor theory (cf. Penrose & MacCallum 1973) and apply it to $A_{\alpha\beta} \in \mathbb{T}_{\alpha\beta}(\mathscr{S})$, as defined by (24). K. P. Tod has tested this for a sphere of symmetry $\mathscr{S}$ in the Schwarzschild, Reissner–Nordstrom and Friedmann–Robertson–Walker (F.R.W.) space–times. He finds, using (27) or (28) with (29), that the Schwarzschild mass is obtained independently of the radius of $\mathscr{S}$, while for the Reissner–Nordstrom space, $m(\mathscr{S})$ becomes reduced at smaller radius by an amount consistent with the fact that less electromagnetic field energy is then incorporated. In the F.R.W. models, $m(\mathscr{S})$ turns out to be precisely equal to the mass that would be surrounded by a sphere in flat space, of surface area equal to that of $\mathscr{S}$, immersed in a fluid of density equal to that of the model. For such models which are spatially 3-spheres, $m(\mathscr{S})$ reaches a maximum when $\mathscr{S}$ is an 'equatorial' 2-sphere and reduces down to zero on either side, consistent with the frequent assertion that the total mass of a closed universe is indeed zero. If the (less plausible) definition (30) with (28) is used, then $m(\mathscr{S})$ is infinite for a maximal equatorial 2-sphere in a F.R.W. model and also at the black-hole horizon in the Schwarzschild solution. This arises, however, because the norm degenerates, which may be regarded as an unreasonable feature of the definition (30).

At null infinity

The conformal invariance of (22), and thus of the definition of the space $\mathbb{T}^{\alpha}(\mathscr{S})$, enables these constructions to be carried out directly at $\mathscr{I}^+$, with a metric that is regular at $\mathscr{I}^+$ and conformal to the physical one. For this purpose it is convenient

to rescale the dyad by $\hat{o}^A = \Omega^{-1} o^A$, $\hat{\iota}^A = \iota^A$, rather than to adopt the symmetrical scaling suggested earlier, where ι^A points along $\mathscr{I}^+$ at $\mathscr{I}^+$. Then $\hat{\omega}^0 = \Omega\omega^0$ and $\hat{\omega}^1 = \omega^1$ attain finite limiting values at $\mathscr{I}^+$, where $\Omega = 0$. In what follows the circumflexes above quantities defined at $\mathscr{I}^+$ are to be understood, not being written in explicitly.

Assume that a standard $\mathscr{I}^+$ exists for $\mathscr{M}$ and is sufficiently smooth that the 4-space curvature for the rescaled metric is C^0 there. We take $\mathscr{M}$ to be vacuum near $\mathscr{I}^+$, or at least with an energy tensor that falls off fast enough that the usual analysis applies. Then we can take Φ_{01}, Φ_{11}, Φ_{21} and Λ as zero, in (23), when $\mathscr{S}$ becomes a cut of $\mathscr{I}^+$. Moreover, we now insert ψ_1, ψ_2, ψ_3 for Ψ_0, Ψ_2, Ψ_3 where the ψ_i are components at $\mathscr{I}^+$ of the rescaled Weyl curvature $\hat{\Psi}_{ABCD} = \Omega^{-1}\Psi_{ABCD}$ (this rescaling preserving the massless spin-2 equations as in (16)) and so $\psi_i = \Psi_i^0$ in the notation of Newman & Penrose (1962). At $\mathscr{I}^+$ we find $\sigma' = 0$, so equations (22) take the simplified form

$$\eth'\omega^0 = 0, \tag{35}$$

$$\eth\omega^1 = \sigma\omega^0, \tag{36}$$

(σ being what is frequently denoted σ^0). We take $\mathscr{I}^+$ to have unit sphere cross-sections; then $\eth$ becomes, in effect, the operator defined by Newman & Penrose (1966). Moreover, the solutions of (35) are the spin-weighted spherical harmonics for $l = \frac{1}{2} = -s$, i.e. linear combinations of ${}_{-\frac{1}{2}}Y_{\frac{1}{2},-\frac{1}{2}}$ and ${}_{-\frac{1}{2}}Y_{\frac{1}{2},\frac{1}{2}}$ and satisfy

$$\eth^2\omega^0 = 0. \tag{37}$$

When $\omega^0 = 0$, the solutions of (36) are written down just as easily, where now the spin-weight s is $+\frac{1}{2}$. These special cases provide Bramson's (1975*a*) 'aligned frames at infinity' and they define a subspace of $\mathbb{T}^\alpha(\mathscr{S})$ to be denoted $I^{\alpha\beta}\mathbb{T}_\beta(\mathscr{S})$. An infinity twistor $I_{\alpha\beta}$ for $\mathbb{T}^\alpha(\mathscr{S})$ can in fact be defined explicitly as follows.

First $I_{\alpha\beta}$ is defined pointwise on $\mathscr{S} \subset \mathscr{I}^+$ as a local twistor, according to equation (5.23) of Penrose & MacCallum (1973):

$$\begin{aligned} I_{\alpha\beta}\underset{1}{Z}{}^\alpha\underset{2}{Z}{}^\beta &= \mathrm{i}\,\underset{1}{\omega}{}^0\underset{2}{\pi}{}_{1'} - \mathrm{i}\,\underset{1}{\omega}{}^0\underset{2}{\pi}{}_{1'} \\ &= \underset{1}{\omega}{}^0\underset{2}{\eth\omega}{}^0 - \underset{1}{\omega}{}^0\underset{2}{\eth\omega}{}^0, \end{aligned} \tag{38}$$

by (32). We find that the $\eth$ derivative of (38) vanishes, from which it follows (because here $s = 0$) that (38) is *constant* over $\mathscr{S}$. Thus, (38) defines $I_{\alpha\beta} \in \mathbb{T}_{[\alpha\beta]}(\mathscr{S})$ with the usual properties; in particular, (38) vanishes if $\underset{1}{Z}{}^\alpha \in I^{\alpha\beta}\mathbb{T}_\beta(\mathscr{S})$.

When $\omega^0 \neq 0$ we can still write down the solutions of (36) explicitly provided we are given one solution of

$$\eth^2\alpha = \sigma \tag{39}$$

(cf. Newman & Penrose 1966). For then (as pointed out to me by K. P. Tod)

$$\omega^1 = \omega^0\eth\alpha - \alpha\,\eth\omega^0 + \xi \tag{40}$$

where ξ has $s = \frac{1}{2}$ and satisfies

$$\eth\xi = 0. \tag{41}$$

We note that, at $\mathscr{I}^+$, $\mathbb{T}(\mathscr{S})$ is *always* four-dimensional since (35) and (41) have two independent solutions each.

 Quasi-local mass and angular momentum in general relativity. *Proc. Roy. Soc. London Ser. A* **381** (1982), no. 1780, 53–63

For a typical (null) component of four-momentum we specialize to $\underset{2}{Z}{}^{\alpha} \in I^{\alpha\beta}\mathbb{T}_{\beta}(\mathscr{S})$ (since $A_{\alpha\beta}I^{\gamma\beta}$ describes the four-momentum) and obtain

$$\zeta = -\frac{\mathrm{i}}{4\pi G}\oint (\psi_2 \underset{1}{\omega^0}\underset{2}{\omega^1} + \psi_3 \underset{1}{\omega^1}\underset{2}{\omega^1})\,\mathscr{S}. \tag{42}$$

We can take $\mathrm{i}\underset{2}{\omega^1} = \bar{\omega}^{0'}$, writing $\underset{1}{Z}{}^{\alpha} = Z^{\alpha}$, since with $\underset{2}{\omega^0} = 0$, (35) and (36) are consequently satisfied ($\eth' = \bar{\eth}$). Then integration by parts yields the Bondi–Sachs expression

$$\zeta = -\frac{1}{4\pi G}\oint |\omega^0|^2(\psi_2 - \sigma N)\,\mathscr{S}, \tag{43}$$

where the news function N satisfies

$$\eth N = \psi_3, \tag{44}$$

(cf. Newman & Penrose 1968). The linearly independent expressions (43) provide four *real* numbers: the components of energy-momentum. (The imaginary parts vanish, as do the three complex components obtained when $\underset{1}{Z}{}^{\alpha}, \underset{2}{Z}{}^{\alpha} \in I^{\alpha\beta}\mathbb{T}_{\beta}(\mathscr{S})$.)

For angular momentum, neither of $\underset{1}{Z}{}^{\alpha}, \underset{2}{Z}{}^{\alpha}$ can lie in $I^{\alpha\beta}\mathbb{T}_{\beta}(\mathscr{S})$ and we obtain, for a typical component (taking $\underset{\mathrm{i}}{Z}{}^{\alpha} = \underset{1}{Z}{}^{\alpha} = Z^{\alpha}$ for simplicity, the quantities $Z^{\alpha}Z^{\beta}$ spanning all of $\mathbb{T}^{(\alpha\beta)}(\mathscr{S})$),

$$\zeta = -\frac{\mathrm{i}}{4\pi G}\oint \{\psi_1(\omega^0)^2 + 2\psi_2\omega^0\omega^1 + \psi_3(\omega^1)^2\}\,\mathscr{S} \tag{45}$$

into which (40) may be substituted. The resulting expression bears some resemblance to previously obtained expressions (cf. Tamburino & Winicour 1966, and more particularly Bramson 1975*b*) but it differs from them by the effective presence of terms involving $\eth$-integrals of σ, e.g. the α of (39). This occurs because the terms in ω^1 cannot all be eliminated by integration by parts and, at best, one is presented with a term

$$2(\psi_2 - N\sigma)\,\omega^0\omega^1,$$

in the integral, into which (40) may be substituted, ω^0 and ξ being finally expressed in terms of $l = \frac{1}{2}$, $s = \mp\frac{1}{2}$ spin-weighted harmonics.

The reason for this significant discrepancy from earlier work is that a quite different attitude is being adopted here with regard to the 'origin' about which the angular momentum is being taken. In the earlier approaches, the cut $\mathscr{S}$ itself would be thought of as providing this 'origin'; or, rather, one would think of a deformed light cone (null hypersurface), meeting $\mathscr{I}^+$ in $\mathscr{S}$, whose 'vertex' (though not actually a point of $\mathscr{M}$) would be the 'origin'. The asymptotic Killing vectors associated with the angular momentum would then be Bondi–Metzner–Sachs (B.M.S.) generators tangential to $\mathscr{S}$. However, an essential feature of the present approach is that the concept of 'origin' becomes largely *decoupled* from the choice of cut. Intuitively, one thinks of a copy of $\mathscr{S}$ lying in the null infinity of a complexified Minkowski space $\mathbb{CM}(\mathscr{S})$ whose 'σ' agrees precisely with that of $\mathscr{S}$. Identifying $\mathscr{S}$ with this

copy, the points of $\mathbb{CM}(\mathscr{S})$ then provide the (complex) four-parameter family of required 'origins'. Only when $\sigma = 0$ (i.e. when $\mathscr{S}$ is a good cut) will this concept of 'origin' agree with the one previously defined, though for *stationary* fields this is inessential, and Bramson (1978) concurs with the present definition. One advantage of the present procedure over the previous ones is that if $\mathscr{S}$ is taken to be a *bad* cut (i.e. $\sigma \neq 0$) of $\mathscr{I}^+$ for $\mathbb{M}$, then the points of $\mathbb{M}$ itself are indeed 'origins', as they should be.

The above space $\mathbb{CM}(\mathscr{S})$ can also be thought of, more constructively, as the complex Minkowski space whose twistor space is $\mathbb{T}^\alpha(\mathscr{S})$ (i.e. whose points are two-dimensional linear subspaces of $\mathbb{T}^\alpha(\mathscr{S})$. To complete the required structure, however, we need a definition $\mathbb{M}(\mathscr{S})$ of *real* points of $\mathbb{CM}(\mathscr{S})$, since the intrinsic spin can always be reduced to zero by moving to a complex origin (Newman 1976, Hansen & Newman 1975). For this (and also for the normalization required in (43), (45)) we need a twistor complex conjugation (26). A problem arises in that neither of the definitions (27) or (28) (with (29) or (30)) is conformally invariant, so that they cannot be immediately applied at $\mathscr{I}^+$. If we use (27) (or equivalently, (28) with (29); but with (30) the norm degenerates completely) we find conformal invariance of $Z^\alpha \bar{Z}_\alpha$ only if $\eth'^2\sigma = \eth^2\sigma'$, which is not generally the case. To circumvent this problem we can choose the asymptotic rest-frame of the Bondi–Sachs four-momentum to provide a *canonical* unit 2-sphere metric on $\mathscr{S}$ and *then* use (27) to define $Z^\alpha \bar{Z}_\alpha$. (This is somewhat inelegant and can perhaps be improved upon.) With this definition, and taking $I^{\alpha\beta} = \overline{I_{\alpha\beta}}$ with $I_{\alpha\beta}$ as in (38), we obtain (25) as required.

The object $A_{\alpha\beta}$ may be thought of as dual to the space of Killing vectors in $\mathbb{M}(\mathscr{S})$. More immediately it is dual to $\mathbb{T}^{(\alpha\beta)}(\mathscr{S})$, which is represented by spinor fields γ^{AB} on $\mathscr{S}$ defined by (20), the ωs satisfying (35), (36). But the condition (25) enables us to pass to a factor space $\mathbb{K}^\alpha_\beta(\mathscr{S})$ of $\mathbb{T}^{(\alpha\beta)}(\mathscr{S})$ defined as the *real* vector subspace of $\mathbb{T}^\alpha_\beta(\mathscr{S})$ consisting of the quantities

$$K^\alpha_\beta = 4\bar{G}_{\beta\gamma} I^{\alpha\gamma} + 4G^{\alpha\gamma} I_{\beta\gamma},$$

where $G^{\alpha\beta} \in \mathbb{T}^{(\alpha\beta)}(\mathscr{S})$ is represented by γ^{AB}. The primary spinor part (the part with two superscripts) of K^α_β is a spinor field $k^{AB'}$ on $\mathscr{S}$ which in $\mathbb{M}$ would be related to γ^{AB} by

$$k^{AA'} = \tfrac{4}{3}(\mathrm{i}\nabla^A_{B'} \bar{\gamma}^{A'B'} - \mathrm{i}\nabla^{A'}_B \gamma^{AB}),$$

where γ^{AB} satisfies (13). This is the spinor version of (12), and k^a is automatically a Killing vector for $\mathbb{M}$. In $\mathscr{M}$, these considerations will apply, instead, to $\mathbb{M}(\mathscr{S})$. When $\mathscr{S}$ is a bad cut none of these 'Killing vectors' is entirely tangential to $\mathscr{S}$, although each one defines a particular B.M.S. generator.

When applied to the Kerr space–time this method gives the mass and spin in their correct ratio (with no anomalous factors of 2) since, in this case, the results are exactly as in the linear theory.

There are intriguing connections between this approach and $\mathscr{H}$-space theory (Newman 1976; Hansen *et al.* 1978; Ko *et al.* 1981). Also there are relations to Witten's (1981) argument for proving positivity of mass, since in each case the

results depend on building things up from the solutions of spinor equations. It should be instructive to follow up these connections. Can a mass positivity theorem be proved at the quasi-local level, for example? Can one obtain an angular momentum flux law?

I am grateful to many colleagues for valuable discussions, and especially to I. M. Singer and K. P. Tod. Further details of this work will be published later.

References

Arnowitt, R., Deser, S. & Misner, C. W. 1962 In *Gravitation* (ed. L. Witten). New York: John Wiley & Sons.

Bondi, H. 1960 *Nature, Lond.* **186**, 535.

Bondi, H., van der Burg, M. G. J. & Metzner, A. W. K. 1962 *Proc. R. Soc. Lond.* A **270**, 103.

Bonnor, W. B. & Rotenberg, M. A. 1966 *Proc. R. Soc. Lond.* A **289**, 247.

Bramson, B. D. 1975*a* *Proc. R. Soc. Lond.* A **341**, 451.

Bramson, B. D. 1975*b* *Proc. R. Soc. Lond.* A **341**, 463.

Bramson, B. D. 1978 *Proc. R. Soc. Lond.* A **364**, 383.

Einstein, A. 1915 *Sber. preuss. Akad. Wiss.* **47**, 844.

Galilei, G. 1638 *Discorsi e dimostrazioni matematiche intorno à due nuoue scienze* (transl. H. Crew & A. deSalvio). New York: Dover (1914).

Geroch, R., Held, A. & Penrose, R. 1973 *J. math. Phys.* **14**, 874.

Geroch, R. & Winicour, J. 1981 *J. math. Phys.* **22**, 803.

Hansen, R. O. & Newman, E. T. 1975 *Gen. Rel. Grav.* **6**, 361.

Hansen, R. O., Newman, E. T., Penrose, R. & Tod, K. P. 1978 *Proc. R. Soc. Lond.* A **363**, 445.

Ko, M., Ludvigsen, M., Newman, E. T. & Tod, K. P. 1981 *Phys. Rep.* **71**, 51.

Lind, R. W., Messmer, J. & Newman, E. T. 1972 *J. math. Phys.* **13**, 1884.

Newman, E. T. 1976 *Gen. Rel. Grav.* **7**, 107.

Newman, E. T. & Penrose, R. 1962 *J. math. Phys.* **3**, 566; **4**, 998.

Newman, E. T. & Penrose, R. 1966 *J. math. Phys.* **7**, 863.

Newman, E. T. & Penrose, R. 1968 *Proc. R. Soc. Lond.* A **305**, 175.

Penrose, R. 1965 *Proc. R. Soc. Lond.* A **284**, 159.

Penrose, R. 1967*a* In *Relativity theory and astrophysics*. 1. *Relativity and cosmology* (ed. J. Ehlers). Providence, Rhode Island: American Mathematical Society.

Penrose, R. 1967*b* *J. math. Phys.* **8**, 345.

Penrose, R. 1968*a* In *Battelles recontres* (ed. C. D. deWitt & J. A. Wheeler). New York: Benjamin.

Penrose, R. 1968*b* *Int. J. theor. Phys.* **1**, 61.

Penrose, R. & MacCallum, M. A. H. 1973 *Phys. Rep.* 6C no. 4, 241.

Penrose, R. & Ward, R. S. 1980 In *General relativity and gravitation* (ed. A. Held), vol. 2. New York: Plenum.

Prior, C. R. 1977 *Proc. R. Soc. Lond.* A **354**, 379.

Sachs, R. K. 1962 *Proc. R. Soc. Lond.* A **270**, 103.

Sachs, R. K. & Bergman P. G. 1958 *Phys. Rev.* **112**, 674.

Streubel, M. 1978 *Gen. Rel. Grav.* **9**, 551.

Tamburino, L. & Winicour, J. 1966 *Phys. Rev.* **150**, 1039.

Winicour, J. 1968 *J. math. Phys.* **9**, 861.

Winicour, J. 1980 In *General relativity and gravitation* (ed. A. Held), vol. 2. New York: Plenum.

Witten, E. 1981 *Communs math. Phys.* **80**, 381.

133
TIME-ASYMMETRY, COSMOLOGICAL UNIFORMITY AND SPACE-TIME SINGULARITIES

Progress in cosmology. Ed. A. W. Wolfendale. Springer, 1982

In my haste at some proof stage or other when I needed a formula in a hurry, I appear to have allowed myself to insert an absurdity on the second page (p. [88]) of this short note, four lines down from the top, where the number $10^{10^3 B^{3/2}}$ appears. This should be $10^{10^{-37}B^2}$.

214 Time-asymmetry, cosmological uniformity and space-time singularities. *Progress in cosmology.* Ed. A. W. Wolfendale, 87–8. 1982

TIME-ASYMMETRY, COSMOLOGICAL UNIFORMITY AND SPACE-TIME SINGULARITIES

Roger Penrose
Mathematical Institute, University of Oxford, U.K.

We know from the "singularity theorems" of general relativity that any universe model which is compatible with the standard interpretation of background radiation and radio source counts, and which satisfies Einstein's equations with physically reasonable assumptions about the positivity of matter density and absence of closed timelike curves, must be singular. Thus the big bang is "stable" in the sense that perturbations away from exact symmetry cannot remove the singularity. Likewise the singularities of gravitational collapse are stable in the same sense. Thus, physics as we presently understand it is limited in its scope and some new physical ideas are needed in order that the regions of the universe referred to as "space-time singularities" can be treated according to the methods of science.

There is, however, some very strong evidence that there are, indeed, powerful laws governing the structure of singularities. This evidence comes from two apparently independent sources. On the one hand there is the observed large-scale uniformity of the universe, which implies that the big-bang singularity was, at least to a high degree of approximation, of a special highly symmetrical type. On the other hand there is indirect evidence which comes from the second law of thermodynamics. The second law in effect implies that the final singularities which arise in collapse must be unconstrained whereas the big bang must be highly constrained. The Bekenstein-Hawking formula for black-hole entropy enables an estimate to be made of the "degree of specialness" of the big bang. In a closed universe containing 10^{80} baryons, for example, this degree of specialness works out as something of the general order of one part in $10^{10^{123}}$.

This suggests that there must indeed have been very precise physical laws in operation at the big bang itself. The new physics involved is necessarily time-asymmetric. A certain "thought-experiment" appears to imply that such time-asymmetric physics should also have relevance to the "observation" problem in quantum mechanics.

The figure 10^{80}, for the baryon number of the universe, is chosen here for convenience only, in order that some calculation can be made. (The popularity of this particular figure seems to stem from the state of the art of cosmology in Eddington's day !) For a universe with B baryons its "specialness" would be about one in $10^{10^3 B^{3/2}}$. This figure is not affected if one supposes the present baryon number B to arise from baryon non-conserving processes in an initially baryon-antibaryon-symmetric universe, since it refers to the possible entropy that can be gained by collapse into black holes, and this entropy easily swamps all other entropy production processes.

While it may be the case that the existence of a neutrino mass of a few tens of eV's will "close the universe", it can be argued that even so, the figure of $B = 10^{80}$ is an "improbable" one - roughly for the reason that this figure leads to a timescale for recollapse which is not far from m_p^{-3} in absolute units ($G = c = \hbar = 1$), but which is related in no simple way to the neutrino mass (the neutrino being taken as the dominant contributor to the mass of the universe). The present apparent universe age is related to the lifetimes of main sequence stars (since it is this that determines "our present epoch"), which leads to a universe lifetime (so far) of the general order of m_p^{-3} in absolute units. In the absence of a theoretical reason that the ultimate recollapse time (if $k = +1$) is also approximately of this order, one is led to guess that the final timescale is $>> m_p^{-3}$ and probably unobservable directly.

References

Carter, B. (1974) In Confrontation of Cosmological Theories with Observational Data (IAU Symp. 63) ed. M.S. Longair. Reidel: Boston.

Penrose, R. (1979) In General Relativity, An Einstein Centenary Survey (ed. S.W. Hawking & W. Israel.) Cambridge University Press.

Penrose, R. (1981) In Quantum Gravity 2: A Second Oxford Symposium (ed. C.J. Isham, R. Penrose & D.W. Sciama) Oxford University Press.

134
REAL-SPACE TOPOLOGY FOR ADVANCED/RETARDED TWISTOR FUNCTIONS

Twistor Newsletter **14** (1982). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston. Longman, 1990

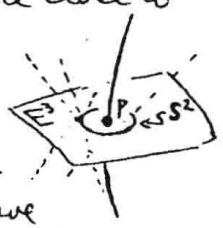
This article and Chapter 135 alternate between two by Toby Bailey where he describes his twistorial treatment of the (anti-self-dual) fields due to charged-particle world-lines. His articles were §1 and §3, while this one is §2 and Chapter 135 is §4. A proper understanding of these two Chapters really requires Toby's two articles. They can be found in *Further Advances in Twistor Theory*, Volume 1 (eds L.J. Mason and L.P. Hughston), Longman, 1990, or *Twistor Newsletter* **14** (1982).

 Real-space topology for advanced/retarded twistor functions. *Twistor Newsletter* **14** (1982). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

§2. Real-space topology for advanced/retarded twistor functions

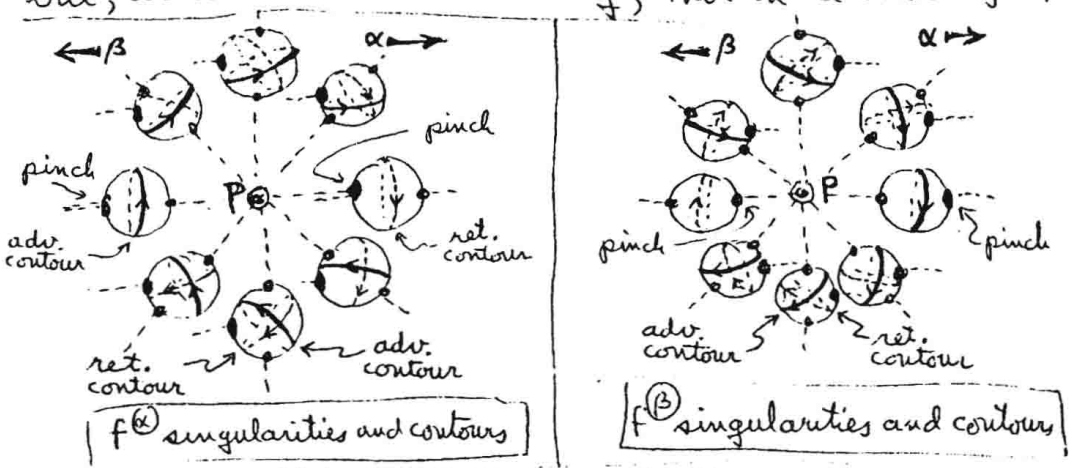
In §§ 1, 3 T.N.B. has shown the importance of the relative cohomology concept for the twistor description of massless fields with source world-lines. The twistor functions that arise from his explicit integral expression (where we assume an analytic timelike curve for the world-line) have some curious and interesting properties. These properties arise in relation to the fact that the twistor functions describe relative cohomology classes in regions in $\mathbb{PT}$. But we can also directly examine the twistor functions in terms of geometry in $\mathbb{M}$, and some useful insights may thereby be gained.

Let us consider points in $\mathbb{M}$ which are close to some point P on the world-line. If they are close enough, the world-line in their neighbourhood may be taken to be almost straight. Taking "time" to be in the direction of the tangent at P to the world-line, we have a "Euclidean" picture in which the source (charge) does not move much from P. Consider a small sphere S^2 centred on P in the 3-plane E^3 orthogonal to this tangent. We shall be concerned with null twistors Z^α representing null lines meeting S^2. We shall picture things in E^3, representing null directions by orthogonal projection into E^3. Null lines in $\mathbb{M}^I$ (and therefore projective null twistors) are thus pictured as points in E^3 (the null-line intersections with E^3) with unit "velocity" vectors attached (describing their null directions).



Take the α- and β-spinors of T.N.B.'s formula to represent directions diametrically opposite in E^3. As in T.N.B.'s article § 1 we take $f^{(\alpha)}$ and $f^{(\beta)}$ to be the twistor functions arising when just the α-pole and just the β-pole, respectively, is surrounded in the s-integration. By examination of the geometry involved

we find that the singularities of $f^{(\alpha)}$, for null twistors incident with a general point Q of S^3, correspond to directions at Q pointing close to P, pointing roughly away from P (these are poles), and pointing in the general direction defined by β. (This last singularity arises because the s-value for the α-pole moves off into the distance (e.g. to infinity) and we no longer are considering only the local geometry of the (complexified) source line near P. In $\mathbb{PT}$, this is the tangency problem encountered by T.N.B.) The singularities for $f^{(\beta)}$ are similar except that now it is the general direction of α that causes problems. Note that at each point Q (in sufficiently general position) we have three singularities in $f^{(\alpha)}$ and thus there are two linearly independent fields that can be computed by contour integration. These are the advanced and retarded fields at Q. Similarly $f^{(\beta)}$ provides both advanced and retarded fields — but, as we shall see shortly, not in the same regions.



In the above figures the singularities and contours for $f^{(\alpha)}$ and $f^{(\beta)}$ are indicated, for various points Q close to P. To obtain Q's on all of S^2, rotate the figures

about the α-β- direction through P. Note that, for $f^{(\alpha)}$, the advanced contour exists everywhere on S except when we get too close to the α-direction from P, whereas the retarded contour fails to exist only when we get too close to the β-direction from P. For $f^{(\beta)}$ the situation is just the other way about. We observe that there is an annular region on S^2 over which $f^{(\alpha)}$ and $f^{(\beta)}$ both give both fields, but each only gives one of them (a different one) on either side.

If we form $f^{(\alpha)} - f^{(\beta)}$, the in and out poles disappear and we are left only with the α and β-singularities. This gives us an "ordinary" cohomology element representing the source-free advanced-minus-retarded wave field.

Let us see that the advanced field is an analytic continuation of minus the retarded field.

We do this by "taking a walk through the complex". Recall that a point of $\mathbb{CM}^I$ with spacelike imaginary part has an S^1's worth of real null lines incident with it (given by the intersection with $\mathbb{PN}$ of the S^2 which is the projective line in $\mathbb{PT}$ representing this point of $\mathbb{CM}^I$). In E^3, such an S^1's worth of null lines appears as a family of unit vectors, one at each point of a circle $\odot$ and such that if $\odot$ is rotated into itself, the unit vector family rotates into itself. A special case is when the unit vectors are all tangent to $\odot$ (with a given orientation). To take our walk, use $f^{(\alpha)}$ and start at the point of S^2 in the α-direction from P, with the depicted retarded contour. Now move back to the left along S^2 spreading this point into a $\odot$ on S^2, with unit vectors tangent to $\odot$ (in accordance with the above "special case"). Keep $\odot$ in planes orthogonal to the α-β-axis and keep its orientation consistent, moving it back over S^2 until it closes up again at the point in the β-direction from P. The contour moves continuously from $+\infty$ to $-\infty$. (Thanks to NB) Pictures missed

135
TWISTOR BLISTERS AND RELATIVE COHOMOLOGY

Twistor Newsletter **14** (1982). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston. Longman, 1990

Here, Toby Bailey's procedures are carried further, by taking into account the observation of Chapter 108 that the twistor space in the presence of electromagnetic charges is in fact *non-Hausdorff*. We find that line bundles over such non-Hausdorff base spaces are described in terms of *relative* cohomology, in accordance with Bailey's descriptions.

 Twistor blisters and relative cohomology. *Twistor Newsletter* **14** (1982). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

§4. Twistor blisters and relative cohomology

T.N.B. shows, in §3, how source multipole-structure shows up as relative sheaf cohomology in twistor space. To get the charge monopole in the anti-self-dual case he needs the sheaf $\mathcal{O}^*$. Now recall (G.A.J.S. & R.P. in TN 9) that the a.-s.-d. Coulomb field could be described as a line bundle over a non-Hausdorff base space consisting of $\mathbb{P}\mathbb{T}$ with the relevant quadric doubled up. From T.N.B.'s description, the a.-s.-d. Coulomb field can also be described as a relative cohomology class. So the question arises as to whether there is some connection between such non-Hausdorff structures and relative cohomology classes. It appears that in fact there is.

First consider T.N.B.'s case. We are concerned with a sequence

[A] $$H^0(\mathcal{N}) \to H^0(\mathcal{N}-\mathcal{R}) \to H^1_{\mathcal{R}} \to H^1(\mathcal{N}) \to H^1(\mathcal{N}-\mathcal{R}) \to \dots$$

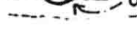
where the sheaf is $\mathcal{O}^*$ throughout, and where $\mathcal{N}$ is a neighbourhood in $\mathbb{P}\mathbb{T}$, of a suitable piece $\mathcal{R}$ of ruled surface in $\mathbb{P}\mathbb{T}$. The required relative cohomology classes, describing the possible source multipole structures for anti-self-dual Maxwell fields, are the elements of $H^1_{\mathcal{R}}$. Now the other four spaces in the sequence [A] are easily interpreted, $H^1(\mathcal{N})$ and $H^1(\mathcal{N}-\mathcal{R})$ describing line bundles over $\mathcal{N}$ and $\mathcal{N}-\mathcal{R}$, respectively. The idea is to interpret $H^1_{\mathcal{R}}$ in terms of line bundles over the non-Hausdorff base space $\hat{\mathcal{N}}$ consisting of $\mathcal{N}$ but with its subspace $\mathcal{R}$ doubled up. We can think of this base space, alternatively, as two copies $\mathcal{N}$ and $\mathcal{N}'$, of $\mathcal{N}$, which are then identified at

corresponding points <u>except</u> at the points $\mathcal{R}$. Thus, the relevant bundles may be thought of as two separate bundles $\mathcal{B}$ and $\mathcal{B}'$, over $\mathcal{N}$ and $\mathcal{N}'$ respectively, which are then identified all over $\mathcal{N}-\mathcal{R}$, but not over $\mathcal{R}$ itself. For this to be possible, of course, the parts of $\mathcal{B}$ and $\mathcal{B}'$ which lie over $\mathcal{N}-\mathcal{R} = \mathcal{N}'-\mathcal{R}$ must be isomorphic as bundles. Now instead of partially identifying the line bundles $\mathcal{B}$ and $\mathcal{B}'$, let us consider them as distinct bundles over $\mathcal{N}$ and <u>divide</u> one by the other. This operation is to provide the map $H^1_{\mathcal{R}} \to H^1(\mathcal{N})$. Indeed, it evidently yields the kernel of the following map $H^1(\mathcal{N}) \to H^1(\mathcal{N}-\mathcal{R})$ since these are the line bundles which trivialize over $\mathcal{N}-\mathcal{R}$, which is precisely the condition that $\mathcal{B}$ and $\mathcal{B}'$ should be isomorphic over $\mathcal{N}-\mathcal{R}$, as required above.

Note that $H^1_{\mathcal{R}}$ does not actually classify line bundles over $\hat{\mathcal{N}}$, but measures only the resultant "twist" between the two branches of the bundle as we pass across $\mathcal{R}$. This is as it should be, because $H^1_{\mathcal{R}}$ only measures the source structure and does not tell us how much radiation is present in the field. The latter information resides in $H^1(\mathcal{N})$, and we would expect to pick that out from the <u>remaining</u> information which resides in the line bundle over $\hat{\mathcal{N}}$. Indeed, this seems to work.

It is also quite easy to see that $\frac{H^0(\mathcal{N}-\mathcal{R})}{H^0(\mathcal{N})}$ provides "twists" between the branches, which are those yielding zero in $H^1_{\mathcal{R}} \to H^1(\mathcal{N})$, consistent with A.

Now this idea seems to have wide application in other contexts, and particularly to the suggestion (cf. R.P. in TN 8) that <u>googlies</u> might be describable by non-Hausdorff structures. In this case it would be the line I (or some blown-up version of I) which might be doubled up in some suitable kind of neighbourhood. For the googly graviton it would not be line bundles but deformations that would concern us. Generally, we would consider replacing $\mathcal{O}^*$ by some sheaf of deformations and it appears, analogously to the above discussion, that the relative cohomology class would measure the "degree of blistering" (say infinitesimally) over the doubled-up subspace. This is just what's needed for the (weak) googly graviton. (Cf. also I.H.B., M.G.E.)



136
ON THE GEOMETRY OF GOOGLY MAPS

Twistor Newsletter **14** (1982). With P. Law. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer). Chapman and Hall/CRC. London, 2001

A study is made of the googly maps in the Eguchi–Hanson complex space-time and a cancellation explained. See Chapter 107 for some non-Hausdorff properties.

 On the geometry of googly maps. *Twistor Newsletter* **14** (1982). With P. Law. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.7 **On the Geometry of Googly Maps** *by R. Penrose and P.R. Law* (TN 14, June 1982)

Law §III.4.5 described the form of the googly maps for the Eguchi-Hanson solution. These have the structure

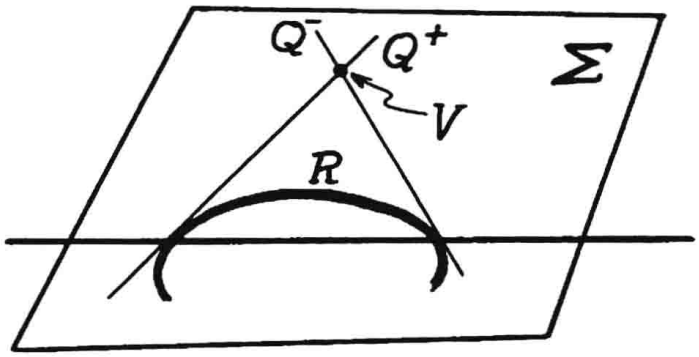
$$\xi^A = \frac{\omega^A R \pm \mathrm{i} Q^{AB}\omega_B R^{1/2}}{Q^+Q^-} \tag{1}$$

where Q^+ and Q^- are quadratic expressions in $Z^\alpha = (\omega^A, \pi_{A'})$ which vanish at $\pi_{A'} = 0$ (so that $Q^\pm = 0$ represent two quadrics through I in $\mathbb{P}T$), with

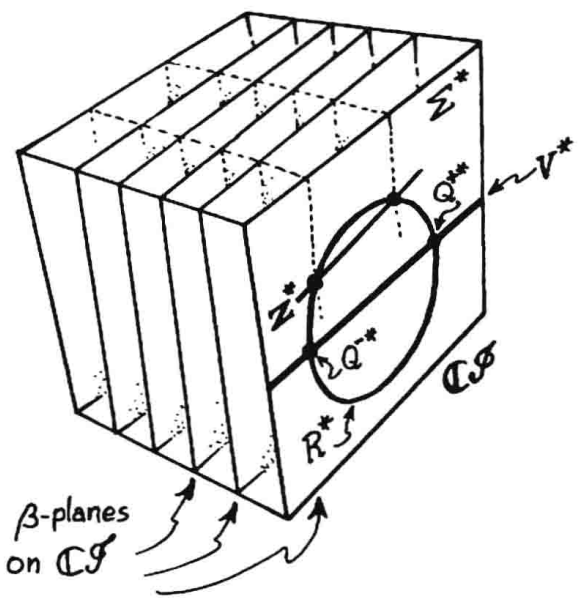
$$Q^+Q^- = Q^{AB}\omega_A\omega_B \quad \text{and} \quad R = Q^+Q^- + L^2M^2,$$

L and M being linear expressions in $\pi_{A'}$. The surface $R = 0$ in $\mathbb{PT}$ is a quartic passing doubly through the line I.

Note that (1) possesses two kinds of singularity, namely, those given by $Q^\pm = 0$ (poles) and those given by $R = 0$. It had at first seemed rather remarkable that if we choose a point on (say) Q^+ which does not lie on Q^-, then on *one* of the branches of (1) the numerator also vanishes, and (1) is regular after all! We shall show how this can be understood in terms of the geometry and indicate that (1) is actually exhibiting important features of general googly maps.



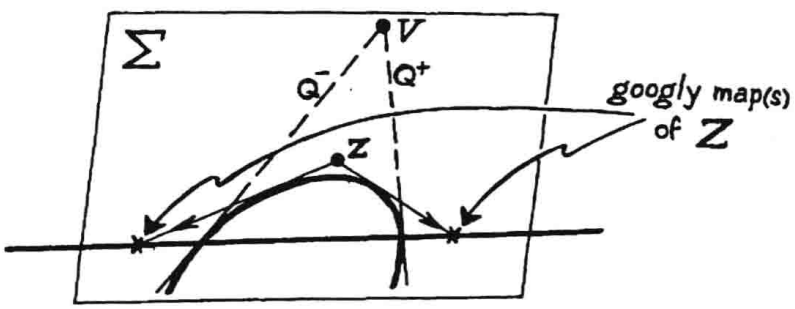
Consider the geometry in some general plane Σ through I (i.e., take $\pi_{0'} : \pi_{1'}$ to be a fixed ratio). The intersections (residual to I) of Σ with Q^+ and $Q^- = 0$ are simply two lines, and with $R = 0$ is a conic which touches these two lines at their intersections with I. Thus $R = 0$ is a branch curve for (1), while $Q^+ = 0$ and $Q^- = 0$ give poles, one on *'each'* of the two branches. To understand this behaviour consider $\mathbb{CP}$. The plane Σ corresponds to a 'dual projective asymptotic twistor entirely on $\mathbb{CP}$' i.e., to a β-plane Σ^* lying on $\mathbb{CP}$. The asymptotic twistor lines in Σ^* represent the points of Σ and, dually, the points of Σ^* represent lines of Σ.



The conic R (shorthand for $\Sigma \cap \{$projective Z's $|$ satisfying $R = 0\}$, etc.) dualizes

 On the geometry of googly maps. *Twistor Newsletter* **14** (1982). With P. Law. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

to a conic R^* in Σ^*, this being the section, by Σ^*, of the good cut representing the point of the Eguchi-Hansom space whose googly map we are describing. Take a twistor line Z^* in Σ^*. This meets R^* in two points (in general) and the resulting (ambiguous) choice of α-plane on $\mathbb{CS}$ through one or the other of the points provides (projectively) the googly map (1) — the ambiguity being reflected in the '$\pm$'. Now the branch surface $R = 0$ corresponds to *contact* between Z^* and R^*, so (as is normal in the reciprocal relation between a point locus and its dual envelope) the *points* of R^* correspond to *tangent lines* to R. Thus the intersections of Z^* with R^* correspond in Σ to the tangent lines from Z to R. The googly maps are obtained geometrically, therefore, by taking the intersections with I of the two tangent lines to R through Z. The pole singularities of (1) arise when the googly map (projectively) becomes one of the points $R \cap I$. It is clear that when Z is on (say) Q^+, but not at $V = Q^+ \cap Q^-$, then only *one* of the choices of googly map is such a point, while the other one remains well-defined — explaining the striking cancellation in (1).



With googly maps in general we expect certain similar features. Again a curve R arises in each Σ (dual to the 'good cut' R^* in Σ^*), and again the googly maps arise in the same way from tangent lines from Z. *Locally* we expect quadratic branching along R in general (though poles like Q^+, Q^- are exceptional).

137
A PUZZLE FOR TWISTORIANS

Twistor Newsletter **14** (1982)

I don't think that the solution was ever provided in $\mathbb{TN}$, but have a look at the introductory remarks to Chapter 211.

A puzzle for twistorians (computer-aided solutions accepted!)

A set $\mathcal{S}$ is equipped with a ternary operation $\mathcal{S}\times\mathcal{S}\times\mathcal{S}\to\mathcal{S}$, denoted $a,b,c \mapsto abc$, subject to (1) $aab = b$, (2) $abb = a$, (3) $(abc)de = ab(cde)$. for all $a, b, \ldots$

PROBLEM A Prove $a(bcd)e = (adc)be$.

PROBLEM B What's this ternary operation all about?

(Hint for problem A: do problem B first!)

138
SOME REMARKS ON GRAVITY AND QUANTUM MECHANICS

Quantum structure of space and time (1981), 3–10. Cambridge University Press, 1982

Exactly as with Chapter 133, there is the same absurd error, this time appearing 13 lines down on the 3rd page of the article (which is p. [5]), where the number $10^{1000B^{3/2}}$ appears. This should be $10^{10^{-37}B^2}$. (I presume that I must have unthinkingly copied this part of one of the articles directly from the other, but I have no recollection of this.)

232 Some remarks on gravity and quantum mechanics. *Quantum structure of space and time* (1981), 3–10, 1982

SOME REMARKS ON GRAVITY AND QUANTUM MECHANICS

Roger Penrose,
Mathematical Institute,
Oxford

The goal of finding an appropriate union of gravitation theory with quantum mechanics has stimulated much fruitful computational activity in recent years. But I think it is also important to keep reviewing the nature of the goals themselves, so as to reduce the likelihood that too much of this activity may be aimed in the wrong direction.

Let us recall a few of the main motivations that have been often stated as reasons for wanting to quantize gravity:

(1) the unity and consistency of physics

(2) the removal of ultraviolet divergences by a gravitational cut-off at the Planck length

(3) the removal, via quantum processes, of the space-time singularities of classical general relativity.

I have no quarrel with (1) as a motivation, except that it is rather non-specific. Some unification between (essentially) the physics of the small (quantum theory) and the large (general relativity) must indeed be part of nature's design. But whether that unification takes the form of a standard quantization procedure applied to standard general relativity theory, or whether the procedure or theory must be replaced by a non-standard one, or whether we should rather be seeking some quite new type of unified scheme having ordinary quantum theory and ordinary general relativity as two distinct limits – all this seems to me to be very far from clear. Perhaps my sympathies lie most with this last suggestion.

As regards (2), it is certainly an appealing possibility that the correct way of incorporating gravity into quantum field theory might remove these divergences. However, we should bear in mind that if (2) is our goal, then it is pointless to look for a renormalizable

theory of quantum gravity. What we should need would necessarily be a finite theory. My own prejudices, however, are that we should be unlikely to find such a finite theory (even with a modified general relativity such as supergravity) so long as we stick with the normal notions of space-time geometry at such small distances. The standard procedures of quantum field theory imply that we do stick with them. (The "space-time foam" calculations of the Cambridge school, for example, I am counting as non-standard.) In fact, the standard (perturbative) procedures are often essentially flat-space ones, with ordinary flat-space causality holding. I shall have more to say about that shortly.

My main quarrel, here, is with (3). I have gradually come around to the view that it is actually misguided to ask that the space-time singularities of classical relativity should disappear when standard techniques of quantum (field) theory are applied to them. Suppose, for the sake of argument, that someone is finally able to show, using some quantization procedure of a reasonably conventional kind, that "space-time singularities" do not occur - or, at least, can be avoided - in quantized gravity. In particular, suppose that the big bang has been eliminated as a singularity in the new scheme, so the universe is seen as arising from a previously collapsing phase which passes through a highly compressed but not quite singular state in order to arrive at its present state of expansion. My personal reaction to this would be one of some disappointment, since we should have lost what seems to me to be the best chance we have of explaining the mystery of the second law of thermodynamics. To explain the second law, we need, primarily, an explanation of why the universe was in such an extraordinarily "special" state in the distant past. Once we have pushed the universe non-singularly back into a previously collapsing phase we are back into the regime of "conventional physics", where there are no rules requiring that the state of the universe be anything particularly "special". A space-time singularity is, almost by definition, "a place where the known laws of physics break down". So it is not unreasonable that principles other than the ones we know and understand should come into play at a singularity. Such a principle ought to have the time-asymmetric character necessary to explain the second law of thermodynamics. Indeed, as I have pointed out elsewhere

(Penrose 1978, 1979, 1981), the "Weyl curvature hypothesis" - according to which initial singularities (TIFs) have to have vanishing Weyl curvature whereas final singularities (TIPs) are unconstrained - appears to achieve this. But if quantum gravity simply removed the singularities, then we should be unable to invoke any new principles and we should be back where we were in our attempts to understand the origin of the second law.

The big bang, whether a singularity or not, was undoubtedly a very highly constrained physical situation. Had there not been any constraining principles (such as the Weyl curvature hypothesis) the Bekenstein-Hawking formula would tell us that the probability of such a "special" geometry arising by chance is at least as small as about one part in $10^{1000B^{3/2}}$ where B is the present baryon number of the universe. If there had been a previous collapsing phase, then this figure would give a measure of how constrained that prior collapsing geometry would have had to have been.

In my own view, whatever it is that we mean by the term "quantum gravity" ought to include an explanation for the grossly time-asymmetrical constraint on whatever physical situation it is that is covered by the term "space-time singularity". In this sense, I believe that the correct theory of quantum gravity must be a time-asymmetric theory.

I should perhaps remark that I do not necessarily believe that B need be finite. (Indeed, the k = -1 cosmological models have a certain mathematical appeal.) An infinite B would not invalidate the "specialness" argument. In fact it would strengthen it! Another point is that even if finite, B may have a value vastly in excess of the popular figure of $B \sim 10^{80}$. It could well be that we are in a very early stage of a universe whose characteristic timescale is enormously greater than the figure of $\sim 10^{10}$ years at which we find ourselves. The arguments provided by Dicke (1962) and Carter (1974) suggesting that our temporal location in the universe should be governed by the timescales of main-sequence stars rather than by the characteristic timescale of the universe as a whole, do have some essential plausibility.

As I have explained elsewhere (Penrose 1981), there seem to be indications that this proposed quantum-gravitational time-asymmetry should, in a different guise, show up in a kind of spontaneous reduction

of the state vector. My own view is that this, also, will emerge as a necessary feature of the correct theory of quantum gravity. It is highly unlikely that any too conventional approach to the quantization of gravity should be able to incorporate these features.

Just to stress my own position on these points, I should mention that the programme of twistor theory seems to be no better placed for producing these desired effects than other programmes. Perhaps there is the general point that twistor theory provides an asymmetrical picture of physics. But this asymmetry is not directly a time-asymmetry; it is rather a parity-asymmetry. Any time-asymmetry seems to have to arise more indirectly.

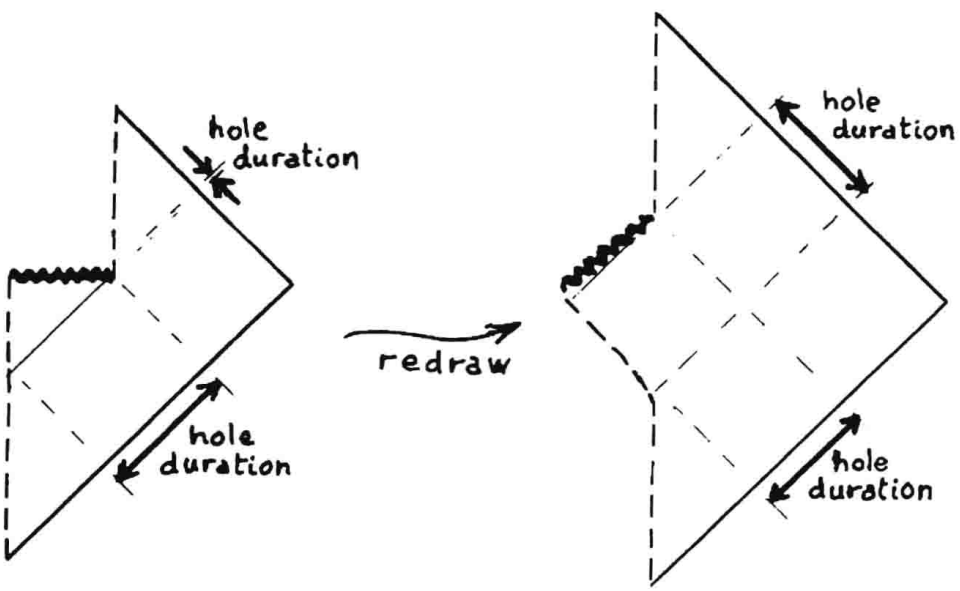
I have, however, been tempted towards the goal of finding a simple twistor description of the big bang. For the conformal flatness entailed by the Weyl curvature hypothesis provides that a direct description of the early geometry in twistor terms exists. The Friedmann-Robertson-Walker cosmologies give the appropriate geometry, and a neat twistor description can indeed be given (Penrose & Rindler 1983), the basic new ingredient being an expression for the space-time metric in terms of two infinity twistors.

Perhaps I ought also to state my present position with regard to the question of time-symmetry of the black-hole-evaporation and -formation processes, since this is relevant to the question of the time-symmetry of quantum gravity. Much discussion of this matter has taken place over the past few years. I think that Don Page has largely persuaded me (Page 1981) that the behaviour, with regard to black hole formation, of the contents of a box with perfectly reflecting walls is somewhat closer to being time-symmetric than I had previously thought. Nevertheless, I would still maintain that time-asymmetries ought to be present. This is not (as he seems to be suggesting) because of any belief that time-asymmetric behaviour would arise out of time-symmetric physics, but because the relevant physics in the box is, in my view, actually time-asymmetric (even in the absence of a quantum gravity theory), because the Weyl curvature hypothesis is involved in the singularity structure. This would entail that only black holes and not white holes could be present in the box. It may well be the case that (for a largish box) black hole disappearance by Hawking evaporation is essentially balanced, as Page maintains, by a production

process which resembles its time-reserve. Nevertheless, manifestly classical collapses must at least occasionally take place, and their time-reverses are not physically allowed for a black hole. For a box containing a large enough mass these manifestly classical collapses would actually be the dominant collapse processes.

The strongest argument against the black hole = white hole picture has always been the incompatibility of their space-time geometries. It is worthwhile to redraw the standard conformal space-time diagram describing spherically symmetric classical collapse to a black hole, followed by its Hawking disappearance, in such a way that it looks as time-symmetric as possible (see fig. 1). There tends to be a considerable freedom in these 2-dimensional pictures since the time-scaling for the in and out light rays is arbitrary. In the familiar left-hand diagram, the "hole duration" seems much longer with respect to advanced time than it does with respect to retarded time. But this is merely an artefact of the particular shape of diagram that has been selected (the choice having been made merely to bring out its

Figure 1
Conformal space-time diagrams of spherically symmetric classical collapse to a black hole, followed by its Hawking disappearance



qualitative features as simply and bluntly as possible). We can keep the radial null rays at 45° and readjust the in and out scales so that the hole duration seems about the same with respect to each, and we arrive at the right-hand diagram in which the space-time geometry appears to be almost time-symmetrical.

I think that from the quantitative point of view, the right-hand diagram is less misleading than the left-hand one, indicating that a time-symmetric picture of the entire physical process may well be usually quantitatively accurate (so long as the box sides are so far off that they can be ignored). The regions of space-time which violate the time-symmetry are qualitatively better illustrated by the left-hand diagram, but these normally contribute only a very small proportion of the 4-volume (say) of the entire process, the right-hand diagram illustrating this better in a quantitative way. However, for an astrophysical black hole, these 4-volumes are nevertheless rather large in absolute terms and it is not really legitimate to ignore their effects altogether.

My point is, then, that while under appropriate circumstances time-symmetry may well dominate the gross behaviour of the contents of the box, the time-asymmetric effects should still be important on an absolute rather than a relative scale (and should become relatively more important as the total mass in the box is increased). In the arguments I presented last year (Penrose 1981) at the Oxford Quantum Gravity 2 conference, I required only that there be a large number of hole-formation modes (namely manifestly classical collapses) which are not adequately balanced by evaporation modes. It is not necessary that these formation modes should be the dominant ones. Moreover, on a broad scale the classical collapse (if it proceeds at the appropriate rate) and quantum evaporation may look roughly like time-reverses of one another, as in the right-hand diagram of the figure, and the effects at large distances could be similar. Nevertheless the geometries still are different locally and a detailed examination of the implications of this fact should repay further analysis.

Finally, I should like to put in a strong plug for geometry in the investigation of quantum gravity. Even Steven Weinberg now informs me that, because of certain "no-go" theorems (Weinberg

& Witten 1980), he is no longer convinced that the anti-geometrical viewpoint is necessarily the most fruitful one! In this connection I should like to point out a result I obtained some years ago, and which is now published in Abe Taub's retirement volume (Penrose 1980). This states that even in the Schwarzschild space-time, with any matter distribution whatever covering the spatially bounded source region, it is not possible to introduce a Minkowskian background metric satisfying the following two requirements: (a) light propagation in the Schwarzschild metric does not violate the background causality, and (b) light rays for the two metrics agree at infinity. The inference to be drawn from this is that perturbation methods, using expansions about Minkowski space in which one attempts to build up a gravitational quantum field theory using the usual flat-space procedures (e.g. "field operators at spacelike-separated points commute") must necessarily break down, presumably because of divergences, even in the simplest cases. It seems to me that approaches to quantum gravity which do not fully take account of curved-space general-relativistic causality will have no chance of success - and this warning applies as much to twistor theorists as to anyone else!

References

Carter, B. (1974) in Confrontation of Cosmological Theories with Observational Data (ed. M.S. Longair; Reidel Publ., Dordrecht).

Dicke, R.H. (1962) Nature (Lond.) 192, 440 .

Page, D. (1981) "Black hole formation in a box", Gen. Rel. Grav. (to appear)

Penrose, R. (1978) in Theoretical Principles in Astrophysics and Relativity (eds. N.R. Lebovitz, W.H. Reid & P.O. Vandervoort; Univ. of Chicago Press).

Penrose, R. (1979) in General Relativity, an Einstein Centenary Survey (eds. S.W. Hawking & W. Israel; Cambridge Univ. Press).

Penrose, R. (1980) in Essays in General Relativity (ed. F.J. Tipler; Adacemic Press, New York).

Penrose, R. (1981) in Quantum Gravity 2, a Second Oxford Symposium (eds. C.J. Isham, R. Penrose & D.W. Sciama; Clarendon Press, Oxford).

Penrose, R. & Rindler, W. (1983) Spinors, Twistors and Space-Time Structure (to be published, Cambridge Univ. Press). For a preliminary account, see Twistor Newsletter 12 (Mathematical Institute 1981).

Weinberg, S. & Witten, E. (1980) Phys. Lett. 96B, 59.

139
GENERAL-RELATIVISTIC KINEMATICS?

Twistor Newsletter **15** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

This article needs to be read in conjunction with Chapter 130 or 132. The whole issue might be re-examined in the light of current cosmological observations strongly suggesting the presence of a *positive cosmological constant*, since this makes the infinity twistor non-degenerate.

 General-relativistic kinematics? *Twistor Newsletter* **15** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.5 **General-relativistic kinematics??** *by R.Penrose* (TN 15, January 1983)

Recall that the standard flat-space *kinematics* of a relativistic object (in the sense of its energy-momentum-angular-momentum structure, the term 'dynamics' applying only when a specific Hamiltonian—or equivalent—has been provided governing the equations of motion of the structure of the object) can be described by the symmetric *kinematic twistor* $A_{\alpha\beta}$ $(= A_{\beta\alpha})$ subject to

$$A_{\alpha\beta}I^{\beta\gamma} = \overline{A_{\beta\gamma}I^{\alpha\gamma}}. \tag{1}$$

The Hermiticity condition (1) is needed to reduce the number of independent components of $A_{\alpha\beta}$ from 10 complex ones (i.e. 20 real ones) to 10 real ones. The obtaining of an appropriate

analogue of (1) has proved to be a major stumbling block in the interpretation of the '$A_{\alpha\beta}(\mathcal{Y})$' that arises in the general-relativistic theory of 2-surface ('superficial') twistors, this $A_{\alpha\beta}(\mathcal{Y})$ being associated with a spacelike topological 2-sphere $\mathcal{Y}$ in (curved) space-time and being intended to describe the 'kinematics' (in the above sense) of the (total) object surrounded by $\mathcal{Y}$ (cf. Penrose in §II.4.2 and Penrose (1982)).

I have always regarded (1) (or its equivalent) as a rather unpleasant looking relation, from the point of view of twistor theory, for such a basic system of physical quantities as that described by $A_{\alpha\beta}$. Perhaps there is some deep-seated connection between this awkwardness and the difficulties in obtaining an appropriate analogue of (1) in general relativity, for a finite 2-surface $\mathcal{Y}$. What I have in mind is the possibility that the 2-surface twistor space $\mathbb{T}^\alpha(\mathcal{Y})$ may have a structure slightly different from that of the standard $\mathbb{T}^\alpha$ associated with Minkowski space $\mathbb{M}$, and that the $A_{\alpha\beta}$ in $\mathbb{M}$ may be describing an *infinitesimal change* away from this standard Minkowskian twistor space structure. The idea is that the $A_{\alpha\beta}(\mathcal{Y})$, for a finite $\mathcal{Y}$ in a curved $\mathcal{M}$, would have a somewhat different structure from the $A_{\alpha\beta}$ in $\mathbb{M}$ and would perhaps describe a *finite* change away from the standard Minkowskian twistor space structure.

How can we associate the Minkowskian $A_{\alpha\beta}$ with an infinitesimal change in twistor space structure? A clue may be found in the 'kinematic sequence' (cf. Penrose & Rindler, 1986, p. 198, and Hughston & Hurd in §II.4.3) where it is found that the space of $A_{\beta\alpha}$'s plays a dual role, either as a subspace of $\mathbb{T}_{(\alpha\beta)}$ (subject to 1), or as a factor space of the trace-free Hermitian part $\hat{\mathbb{H}}^\alpha_\beta$ of $\mathbb{T}^\alpha_\beta$. We can regard $\hat{\mathbb{H}}^\alpha_\beta$ as the space of real conformal Killing vectors k^a in $\mathbb{M}$, these providing infinitesimal motions of $\mathbb{T}^\alpha$ preserving the twistor norm $Z^\alpha\bar{Z}_\alpha$. The relation between k^a and the corresponding angular momentum structure, given by M^{ab} regarded as a field (a function of the origin point), may be expressed as

$$M^{ab} = \nabla^{[a}k^{b]} \tag{2}$$

(cf. Penrose & MacCallum 1973 and Penrose & Rindler 1986). The required dependence on position of M^{ab} is a consequence of (2) (where $\nabla^{(a}k^{b)} = \frac{1}{4}g^{ab}\nabla_c k^c$) and all M^{ab}'s arise in this way. The translations and dilations (and their compositions) are the k^a's which are killed by (2) so we may regard the space of M^{ab}-fields (i.e. the space of $A_{\alpha\beta}$'s) as the space of conformal Killing vectors for $\mathbb{M}$ *factored out* by the translations and dilations. Thus to describe this space (or its non-linear 'exponentiation') we seek some structure in $\mathbb{M}$ which is invariant precisely under these translations and dilations and look for the family of conformal transforms (infinitesimal or finite)

of this structure.

One such structure is the 'projective π-projection' of $\mathbb{T}$:

$$\mathbb{T}^\alpha \to \{\mathbf{S}_{A'}\} \tag{3}$$

where $\{\mathbf{S}_{A'}\}$ stands for the *projective* space of $\pi_{A'}$'s, taken as a *given* space. Thus a proposal for a 'non-linear $A_{\alpha\beta}$' would be a map (3) which is a general $SU(2,2)$-transform of the standard one. Geometrically, (2) is described by a line $\hat{I}$ in $\mathbb{PN}$, where the family of planes through $\hat{I}$ is given a parameterization as a *standard* Riemann sphere $\mathbb{CP}^1$ (namely $\{\mathbf{S}_{A'}\}$), the fibres of the projection being the individual planes. It is perhaps a little easier to think in dual terms, where the injection

$$\{\mathbf{S}^A\} \to \{\mathbb{T}^\alpha\} \tag{4}$$

plants a standard Riemann sphere $\mathbb{CP}^1$ as a parameterized projective line $\hat{I}$ in $\mathbb{PN}$, and the different ways of doing this give the required space. Note that there is a 10-parameter (real) freedom in doing this (4 for the line $\hat{I}$; 6 for the different choices of parameterization on $\hat{I}$). When $\hat{I}$ lies infinitesimally separated from I, with an infinitesimal shift in parameterization, we recover the original space of 'linear' $A_{\alpha\beta}$'s.

I am trying to put forward a minor modification of this tentative suggestion. In place of (3) and (4), consider the possible exact sequences

$$0 \to \mathbf{S}^A \to \mathbb{T}^\alpha \to \mathbf{S}_{A'} \to 0 \tag{5}$$

where $\mathbf{S}^A$ is a *given* 2-dimensional complex vector space and $\mathbf{S}_{A'}$ is its complex conjugate dual. $\mathbb{T}^\alpha$ is also taken to be a *given* space, namely a particular 4-dimensional complex vector space with standard $(++--)$ twistor norm $Z^\alpha \bar{Z}_\alpha$, but the maps can vary. (5) is to be compatible with this structure of $\mathbb{T}^\alpha$ in the (standard) sense that the dual sequence

$$0 \leftarrow \mathbf{S}_A \leftarrow \mathbf{T}_\alpha \leftarrow \mathbf{S}^{A'} \leftarrow 0 \tag{6}$$

to (5) is also the complex conjugate of (5) (in the reverse order). In terms of indexed quantities, the sequences (5), (6) can be expressed by use of objects

$$e_{A'\alpha}\,, \quad e_A{}^{\alpha} \tag{7}$$

where

$$\overline{e_{A'\alpha}} = e_A{}^{\alpha} \tag{8}$$

and

$$e_{A'\alpha}\, e_B{}^{\alpha} = 0, \tag{9}$$

$\mathbf{S}^A \to \mathbb{T}^\alpha$ being achieved by $\omega^A e_A{}^\alpha = Z^\alpha$ and $\mathbb{T}^\alpha \to \mathbf{S}_{A'}$ by $Z^\alpha e_{A'\alpha} = \pi_{A'}$, etc.

Suppose we have two such pairs of objects, namely (7), subject to (8), (9) and

$$f_{A'\alpha}\,, \quad \text{with } f_A{}^{\alpha} = \overline{f_{A'\alpha}}\,, \quad f_{A'\alpha} f_B{}^{\alpha} = 0. \tag{10}$$

Let us define

$$A_{\alpha\beta} = -iCe_{A'(\alpha} f^{A'}{}_{\beta)} \tag{11}$$

for some positive constant C, and

$$B_{\alpha\beta} = e_{A'[\alpha} f^{A'}{}_{\beta]} \tag{12}$$

(index raised with the 'standard' $\varepsilon^{A'B'}$ of $\mathbf{S}_{A'}$). When $f_{A'\alpha} = e_{A'\alpha}$ = the *canonical* $\pi_{A'}$-projection for $\mathbb{M}$, then $A_{\alpha\beta} = 0$ and $B_{\alpha\beta} = I_{\alpha\beta}$. When $f_{A'\alpha}$ differs from the canonical $e_{A'\alpha}$ by a small quantity of order ε, then $A_{\alpha\beta}$ is of order ε and is, to 1st order in ε, a quantity having the structure of a kinematic twistor (satisfying (1) to first order in ε), where $I_{\alpha\beta} = B_{\alpha\beta} + 0(\varepsilon)$. For a general $A_{\alpha\beta}$, $B_{\alpha\beta}$ given by (11), (12) we have the Hermiticity property

$$A_{\alpha\gamma}\overline{B}^{\beta\gamma} = \overline{A}^{\beta\gamma} B_{\alpha\gamma}, \tag{13}$$

like (1), but $B_{\alpha\beta}$ is not normally 'simple'. (Certain restrictions may be put on $e_{A'\alpha}$, $f_{A'\alpha}$ if $\varepsilon_{\alpha\beta\gamma\delta}$ is regarded as part of the structure of $\mathbb{T}^\alpha$ e.g.

$$e_{A'[\alpha}\, e^{A'}{}_{\beta]} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}\, e_A{}^{\gamma} e^{A\delta}, \tag{14}$$

and the corresponding relation for the fs, but the status of these is unclear.)

Recall now, the Tod form of the expression for $A_{\alpha\beta}(\mathcal{Y})$:

 General-relativistic kinematics? *Twistor Newsletter* **15** (1983). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. 1995

$$A_{\alpha\beta}(\mathcal{S}) = -\frac{i}{4\pi G}\oint (\pi_{\alpha 0'}\pi_{\beta 1'} + \pi_{\alpha 1'}\pi_{\beta 0'})\mathcal{S} \tag{15}$$

($\mathcal{S}$ = surface-area 2-form on $\mathcal{S}$) (cf. Penrose in §II.4.2 and Penrose 1982), and also the local expression on $\mathcal{S}$

$$\pi_{\alpha 0'}\pi_{\beta 1'} - \pi_{\alpha 1'}\pi_{\beta 0'} \tag{16}$$

for $I_{\alpha\beta}$. This tantalizingly suggests a relation between (15) and (11), and between (16) and (12), with something like an integral of $\pi_{\alpha 0'}\pi_{\beta 1'}$ over $\mathcal{S}$ producing the quantity

$$Se_{A'\alpha}f^{A'}{}_{\beta} = SB_{\alpha\beta} + \frac{iS}{C}A_{\alpha\beta} \tag{17}$$

(S being the total surface area of $\mathcal{S}$, and $C = \frac{S}{4\pi G}$). For this to make sense, the integral of $\pi_{\alpha 0'}\pi_{\beta 1'}$ would have to have matrix rank 2 (as an element of $\mathbb{T}_{\alpha\beta}$). I have not been able to prove (or disprove) this, and the whole scheme remains highly conjectural.

References

Penrose, R. (1982) Quasi-local mass and angular momentum in general relativity, Proc. Roy. Soc. **A381**, 53.

Penrose, R. & MacCallum, M.A.H. (1973) Twistor theory: an approach to quantization of fields and space-time, Phys. Rep. **6C**, 241-315

Penrose, R. & Rindler W. (1986) *Spinors and space-time*, Vol. 2 , C.U.P.

140
SPINORS AND TORSION IN GENERAL RELATIVITY

Foundations of Physics **13** (1983), no. 3, 325–40

I have always found the issue of torsion in general relativity somewhat enigmatic, and I had hoped that the considerations of this paper might shed some light on the ECSK theory. Unfortunately, I still find it enigmatic, especially since a colleague has shed some doubt on the complete accuracy of my equations on p. [333], where a sign error (or something) in (49) forbids the deduction of (53). Unfortunately, I have not yet found the time to check all this.

Foundations of Physics, Vol. 13, No. 3, 1983

Spinors and Torsion in General Relativity

Roger Penrose[1]

Received August 18, 1982

Conformal rescalings of spinors are considered, in which the factor Ω, in $\varepsilon_{AB} \mapsto \Omega\varepsilon_{AB}$, is allowed to be complex. It is argued that such rescalings naturally lead to the presence of torsion in the space-time derivative ∇_a. It is further shown that, in standard general relativity, a circularly polarized gravitational wave produces a (nonlocal) rotation effect along rays intersecting it similar to, and apparently consistent with, the local torsion of the Einstein–Cartan–Sciama–Kibble theory. The results of these deliberations are suggestive rather than conclusive.

1. INTRODUCTION

It is an especial pleasure for me to have this opportunity to pay my respects to Paul Dirac. There is, of course, the obvious pleasure of honoring a great physicist, but also there is a more personal one, in that it was he who taught me not only quantum mechanics but also the formalism of 2-component spinors, in lectures, at Cambridge University, of superb clarity and elegant precision. Such spinors have been a major influence in my own research over some 25 years.

The first indications of the fundamental importance to physics of the (Lorentzian) spinor concept were, of course, presented in Dirac's 1928 epoch-making papers.[(1)] But it is my opinion that we have still not yet seen the full significance of spinors—particularly the 2-component ones—in the basic structure of physical laws.

The questions I shall discuss here concern the role of *torsion* in general relativity. I shall present two arguments, one arising out of the mathematics of spinors, and the other more directly physical, which have a bearing on this matter, though both are suggestive only and do not have unambiguous

[1] Mathematical Institute, Oxford, England.

implications with regard to the physical role of torsion. The second argument specifically relates to the so-called Einstein–Cartan–Sciama–Kibble (ECSK) theory (cf. Cartan,[2] Sciama,[3] Kibble,[4] Trautman,[5] Hehl *et al.*[6]). The first might or might not have to do with that theory, but it seemingly has implications also for any theory which finds natural representation in 2-component spinor terms.

2. COMPLEX CONFORMAL RESCALINGS

Consider a conformal rescaling

$$g_{ab} \mapsto \hat{g}_{ab} = \Omega^2 g_{ab} \tag{1}$$

of the space-time metric g_{ab} (signature $+---$) where Ω is a twice-differentiable *real* scalar function on the space-time $\mathscr{M}$. Adopting standard abstract-index conventions and 2-spinor notation (cf. Penrose,[23] Penrose and Rindler[8]) we have

$$g_{ab} = \varepsilon_{AB}\varepsilon_{A'B'} \tag{2}$$

and it is usual to accompany (1) by

$$\varepsilon_{AB} \mapsto \hat{\varepsilon}_{AB} = \Omega\varepsilon_{AB}, \quad \varepsilon_{A'B'} \mapsto \hat{\varepsilon}_{A'B'} = \Omega\varepsilon_{A'B'} \tag{3}$$

There is, however, a generalization of (3) which is very suggestive. This is experessed by allowing Ω to become *complex* and writing the more general

$$\varepsilon_{AB} \mapsto \hat{\varepsilon}_{AB} = \Omega\varepsilon_{AB}, \quad \varepsilon_{A'B'} \mapsto \hat{\varepsilon}_{A'B'} = \bar{\Omega}\varepsilon_{A'B'} \tag{4}$$

in place of (3), so that, by (2), (1) becomes

$$g_{ab} \mapsto \hat{g}_{ab} = \Omega\bar{\Omega}g_{ab} \tag{5}$$

The particular situation in which Ω is taken to have unit modulus everywhere:

$$\Omega = e^{i\theta} \ (\theta \text{ real}) \tag{6}$$

was considered by Infeld and van der Waerden[9] in their original paper, the intention being to interpret the phase changes induced by (4) as those which generate the electromagnetic field. However, that particular interpretation proved inappropriate, partly because it implied a definite relationship between the spin and charge of a particle, a relationship that had just been refuted by the discovery of the neutron in 1932. It may yet be the case that

the transformations (4), with (6), will prove to have deep connections with electromagnetism, but for the moment I propose to persue a different line of thought.

The equations of physics which are best known to be conformally invariant are those for massless free fields (cf. Dirac,[10] Fierz[11]):

$$\nabla^{AA'}\phi_{AB\ldots L} = 0 \tag{7}$$

where $\nabla_a = \nabla_{AA'}$ is the standard Christoffel (=Levi Civita) differentiation operator for the space-time $\mathscr{M}$, here extended so that it applies to spinors, and uniquely characterized by the fact that

$$\nabla_a \varepsilon_{BC} = 0, \qquad \nabla_a \varepsilon_{B'C'} = 0 \tag{8}$$

together with the vanishing of the torsion tensor $T_{ab}{}^{c}$:

$$T_{ab}{}^{c} = 0 \tag{9}$$

Here the torsion is defined by

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) f = T_{ab}{}^{c} \nabla_c f \tag{10}$$

(for arbitrary scalar f); and the normal additivity and Leibniz properties

$$\begin{aligned} \nabla_a(\psi_\ldots + \chi_\ldots) &= \nabla_a \psi_\ldots + \nabla_a \chi_\ldots \\ \nabla_a(\psi_\ldots \chi_\ldots) &= \psi_\ldots \nabla_a \chi_\ldots + (\nabla_a \psi_\ldots)\chi_\ldots \end{aligned} \tag{11}$$

are being assumed for ∇_a (cf., for example, Penrose and Rindler[8]).

The field $\phi_{A\ldots L}$ is assumed symmetric in its indices:

$$\phi_{AB\ldots L} = \phi_{(AB\ldots L)} \tag{12}$$

(where round brackets denote symmetrization and square ones anti-symmetrization). The number, n, of indices of $\phi_{AB\ldots L}$ is twice the spin of the field. When $n = 1$, (7) becomes the Dirac–Weyl equation for the neutrino (cf. Weyl[12]) and when $n = 2$, it becomes the spinor form (cf. Laporte and Uhlenbeck[13]) of the Maxwell free-field equations—or rather their anti-self-dual part, in the case of complex fields. If we regard $\phi_{AB\ldots L}$ as a (positive frequency) complex wave function, then solutions of (7) provide us with *left-handed* helicity states. To obtain the *right-handed* helicity states as well, we need to pair $\phi_{A\ldots L}$ with another symmetric spinor $\chi_{A'\ldots L'}$, of complex-conjugate type, satisfying

$$\nabla^{AA'}\chi_{A'B'\ldots L'} = 0 \tag{13}$$

In (conformally) flat $\mathscr{M}$, Eqs. (7) and (13) are satisfactory for all spins ($n \neq 0$), but in a general curved $\mathscr{M}$ they encounter the familiar consistency problems if $n \geqslant 3$ (Buchdahl,[14] Plebanski[15]). In standard classical general relativity, with $n = 4$ and $\phi_{ABCD} = \Psi_{ABCD}$, where Ψ_{ABCD} is the Weyl curvature spinor (cf. Penrose,[16] Penrose and Rindler[8]), these problems disappear, as they do also with $n = 3$ for the graviton's partner in supergravity theory (cf. Deser and Zumino[17])—although for that theory the torsion (10) is non-vanishing.

Under the conformal rescaling (1) and (3), for real Ω, conformal invariance for (7) is achieved by taking $\phi_{A\ldots L}$ to have the appropriate conformal weight.

$$\phi_{A\ldots L} \mapsto \hat{\phi}_{A\ldots L} = \Omega^{-1}\phi_{A\ldots L} \tag{14}$$

(cf. Cunningham,[18] Bateman,[19] Dirac,[20] McLennan,[21] and Penrose[22,23]). The invariance of (7), with (14) is immediately obtainable once we have the general formula for transforming the Christoffel derivative under (1) (cf. Buchdahl,[24] Penrose[22]):

$$\begin{aligned}\hat{\nabla}_{AA'}\psi^{B\ldots H'}_{E\ldots K'} = \nabla_{AA'}\psi^{B\ldots H'}_{E\ldots K'} - Y_{EA'}\psi^{B\ldots H'}_{A\ldots K'} - \cdots - \bar{Y}_{K'A}\psi^{B\ldots H'}_{E\ldots A'} \\ + \varepsilon_A{}^B Y_{B_0A'}\psi^{B_0\ldots H'}_{E\ldots K'} + \cdots + \varepsilon_{A'}{}^{H'}\bar{Y}_{H_0'A}\psi^{A\ldots H_0'}_{E\ldots K}\end{aligned} \tag{15}$$

where

$$Y_a = \Omega^{-1}\nabla_a\Omega = \nabla_a \log \Omega \tag{16}$$

and where the use of bars (unnecessary here since Y_a is real) is for later convenience.

In a (conformally) flat $\mathscr{M}$, there is a conformally invariant form of potential for $\phi_{A\ldots L}$ which is of some significance (cf. Penrose and Ward 1980,[25] Eastwood, Penrose and Wells[26]):

$$\phi_{AB\ldots L} = \nabla^{B'}_{(B\ldots}\nabla^{L'}_{L}\theta_{A)B'\ldots L'} \tag{17}$$

where

$$\nabla^{A(A'}\theta^{B'\ldots L')}_{A} = 0 \tag{18}$$

$$\theta^{B'\ldots L'}_{A} = \theta^{(B'\ldots L')}_{A} \tag{19}$$

the conformal invariance being achieved by

$$\hat{\theta}^{B'\ldots L'}_{A} = \Omega^{-1}\theta^{B'\ldots L'}_{A} \tag{20}$$

Also conformally invariant is the *twistor* equation

$$\nabla^{(A}_{A'}\omega^{B)} = 0 \tag{21}$$

whose invariance is achieved by

$$\hat{\omega}^A = \omega^A \tag{22}$$

We now ask what happens if we replace (1) and (3), for real Ω, by (5) and (4), for complex Ω. After rescaling the ε's we can obtain the new Christoffel derivative ∇_a, defined uniquely by the "hatted" versions of (8), (9), and (10). But now we find (cf. Eastwood, Penrose, and Wells[26]) that (15) is replaced by a similar expression but with $Y_{AA'}$ (and $\bar{Y}_{A'A}$) replaced by

$$\tfrac{1}{2}(Y_{AA'} + \bar{Y}_{A'A}) \tag{23}$$

and with an additional term

$$\tfrac{1}{4}r(Y_{AA'} - \bar{Y}_{A'A})\psi^{B\ldots H'}_{E\ldots K'} \tag{24}$$

added on the right, where r is the number of upper unprimed plus lower primed minus upper primed minus lower unprimed indices of $\psi^{\cdots}_{\cdots}$. We assume (16), as before, but now Y_a is complex.

With the additional complications resulting from the terms (23) and (24) we find that conformal invariance of (7) and of (18) is achieved if the factor Ω^{-1} is replaced by

$$\Omega^{\frac{n}{4}-\frac{1}{2}}\bar{\Omega}^{-\frac{n}{4}-\frac{1}{2}} \tag{25}$$

and (22) is replaced by

$$\hat{\omega}^A = \Omega^{-1/4}\bar{\Omega}^{1/4}\,\omega^A \tag{26}$$

These conformal weights seems rather ugly, however. If, instead of modifying (15) with (23) and (24), we were to stick to the original formula (15), we should find (not surprisingly—since the original calculation would now be unchanged) that the original much more pleasant scaling behavior (14), (20), and (22) would be retained!

So what is the objection to using the ∇_a defined by (15) and (16) when Ω is complex? The necessary properties (8) and (11) still hold, but now instead of (9) we have a *torsion* whose value is

$$\hat{T}_{ab}{}^{c} = i(Y_d - \bar{Y}_d)e_{ab}{}^{cd} \tag{27}$$

where e_{abcd} is the Levi Civita tensor (with $e_{abcd} = e_{[abcd]}$ and $e_{0123} = 1$ in a right-handed orthonormal frame). More generally, if we allow the original ∇_a to have a nonvanishing torsion $T_{ab}{}^c$ then we obtain

$$\hat{T}_{ab}{}^c = T_{ab}{}^c + i(Y_d - \bar{Y}_d)e_{ab}{}^{cd} \tag{28}$$

It seems to me that if we are going to allow the complex rescalings (4), then we ought also to bring in the torsion defined by (27), or to modify an already existing torsion by (28). The conformal invariance properties that I have mentioned appear to be important, and it seems a pity to have to introduce the ugly factors of (25) and (26). Moreover, the change in the torsion (28) supplies, apart from an overall constant phase factor, the otherwise missing information in the complex rescaling Ω.

3. REPRESENTATION OF A 2-SPINOR

Recall the geometrical representation (up to sign) of a 2-spinor κ^A as a *null flag* (cf. Payne[27] and Penrose[23]). The flagpole is the future-null vector

$$k^a = \kappa^A \bar{\kappa}^{A'} \tag{29}$$

and the flag plane is defined by the null bivector

$$K^{ab} = \kappa^A \kappa^B \varepsilon^{A'B'} + \varepsilon^{AB} \bar{\kappa}^{A'} \bar{\kappa}^{B'} \tag{30}$$

If we assume that κ^A is unaltered by the transformation (4) (like ω^A, in (22)) then we find that the flagpole k^a is also unaltered, but that

$$K^{ab} \mapsto W^{-1}(K^{ab} \cos\theta + {}^*K^{ab} \sin\theta) \tag{31}$$

where the *dual* ${}^*K^{ab}$ of K^{ab} is

$${}^*K^{ab} = \tfrac{1}{2} K_{cd} e^{abcd} \tag{32}$$

and

$$\Omega = We^{i\theta} \quad (W > 0, \theta \text{ real}) \tag{33}$$

Thus the flag plane rotates through an angle θ, in a right-handed sense, about the flagpole direction under (4). If, on the other hand, we assume that κ_A is unaltered, i.e., that $\kappa^A \to \Omega^{-1}\kappa^A$, then we find that the flag plane rotates through θ in a left-handed sense.

This is perhaps a useful observation (cf. below) since it tells us how much rotation is induced on a parallelly propagated polarization plane, by

that torsion (27) which results from a complex rescaling (4). For if k^a is tangent to a null geodesic γ, and the flagpole and flag plane are parallelly propagated along γ according to

$$k^a \nabla_a \kappa^B = 0 \tag{34}$$

we find that this is preserved under (15) if

$$\kappa_A \mapsto \kappa_A, \qquad \text{i.e., } \kappa^A \to \Omega^{-1} \kappa^A \tag{35}$$

From the above, we therefore see that if a polarization plane (flag plane) is propagated parallelly according to ∇_a and to $\hat{\nabla}_a$, respectively, along a null geodesic γ from P to Q, then there will be a discrepancy $\theta(P) - \theta(Q)$ in the respective resulting polarization directions. (We note that the definition of a null geodesic is not, however, altered in the passage from ∇_a to $\hat{\nabla}_a$.)

4. CURVATURE

A few words concerning the spinor description of the curvature tensor in the presence of torsion will not be out of place here. We assume that (8) holds, with a standard symmetric Lorentzian metric given by (2), but that a torsion (10) is present, the curvature tensor R_{abcd} being defined by

$$(\nabla_a \nabla_b - \nabla_b \nabla_a - T_{ab}{}^c \nabla_c) V^d = R_{abc}{}^d V^c \tag{36}$$

Here we have

$$R_{abcd} = R_{[ab][cd]} \tag{37}$$

but

$$R_{[abc]d} = -T_{[ab}{}^e T_{c]ed} - \nabla_{[a} T_{bc]d} \tag{38}$$

Owing to the nonvanishing of the right-hand side of (38) we find that R_{abcd} is not generally symmetric under interchange of ab with cd and consequently the Ricci tensor

$$R_{ab} = R_{acb}{}^c \tag{39}$$

is normally not symmetric. In the ECSK theory this nonsymmetric Ricci tensor is related to the (now nonsymmetric) energy tensor E_{ab} by the standard-looking Einstein equation

$$R_{ab} - \tfrac{1}{2} g_{ab} R = -8\pi G E_{ab} \tag{40}$$

where $R = R_a{}^a$ and G is Newton's gravitational constant (although owing to the ordering of indices naturally arising in (36), the index-ordering for R_{ab} and E_{ab} appears to be the reverse of what is usual in the literature). The lack of symmetry in E_{ab} in this theory is due to the presence of intrinsic spin, whose density is described by a tensor

$$S_{ab}{}^c = -S_{ba}{}^c \tag{41}$$

related to the torsion by

$$T_{ab}{}^c + 2g_{[a}{}^c T_{b]d}{}^d = -8\pi G S_{ab}{}^c \tag{42}$$

i.e., by

$$T_{ab}{}^c = -8\pi G(S_{ab}{}^c + g_{[a}{}^c S_{b]d}{}^d) \tag{43}$$

and we have

$$(\nabla_c - T_{dc}{}^d)S_{ab}{}^c = E_{ba} - E_{ab} \tag{44}$$

(A discussion of the spinor form of (42) and (43) is given in Penrose and Rindler[8]).

The number of independent (real) components of R_{abcd} at each point turns out to be 36, which is 16 more than that in standard general relativity. The information of these 36 components is conveniently encoded in quantities

$$\Psi_{ABCD}, \qquad \Phi_{ABC'D'}, \qquad \Sigma_{AB}, \text{ and } \Lambda \tag{45}$$

with symmetries

$$\Psi_{ABCD} = \Psi_{(ABCD)}, \qquad \Phi_{ABC'D'} = \Phi_{(AB)(C'D')}, \qquad \Sigma_{AB} = \Sigma_{(AB)} \tag{46}$$

and having the respective numbers of complex components:

$$5,\ 9,\ 3, \text{ and } 1 \tag{47}$$

Here Ψ_{ABCD} is the Weyl spinor (with conformal invariance as described below), while $\Phi_{ABC'D'}$, Σ_{AB} and Λ provide the information of, respectively, the trace-free symmetric part, antisymmetric part, and trace of the complex "Ricci tensor" constructed from the left-anti-self-dual part

$$\tfrac{1}{2}R_{abcd} + \tfrac{1}{4}i\varepsilon_{ab}{}^{pq}R_{pqcd} \tag{48}$$

of the curvature. The quantity Σ_{AB} and the combinations $\Phi_{ABC'D'} - \bar{\Phi}_{C'D'AB}$, $\Lambda - \bar{\Lambda}$ are determined directly by the torsion and its derivative, so when the

torsion vanishes, $\Sigma_{AB} = 0$, while $\Phi_{ABC'D'}$ and Λ reduce to the standard Ricci "curvature spinors" of general relativity (cf. Penrose,[16] Penrose and Rindler[8]).

The Weyl spinor Ψ_{ABCD} can be defined in the standard way from R_{abcd}. It satisfies

$$\{\nabla_{D'(A}\nabla_{B)}{}^{D'} - \tfrac{1}{2}T_{D'(AB)}{}^{D'P}\}\nabla_{|P|}\kappa_{C)} = \Psi_{ABCD}\kappa^{D} \tag{49}$$

from which its conformal invariance under the complex rescalings (4) (with (15), (16), and (28)) may be derived by the following argument. Assume that κ^A scales like ω^A as in (22), i.e., $\kappa_A = \Omega\kappa_A$, then as in the invariance of (21) we have

$$\hat{\nabla}_{A'(A}\hat{\kappa}_{B)} = \Omega\nabla_{A'(A}\kappa_{B)} \tag{50}$$

Now if $\lambda_{ABC'} = \lambda_{(AB)C'}$ scales as

$$\hat{\lambda}_{ABC'} = \Omega^2\bar{\Omega}^{-1}\lambda_{ABC'} \tag{51}$$

then, as with (18), we have

$$\hat{\nabla}^{C'}_{(A}\hat{\lambda}_{BC)C'} = \Omega^2\bar{\Omega}^{-2}\nabla^{C'}_{(A}\lambda_{BC)C'} \tag{52}$$

Equating (50) with (51) and applying (52), where the discrepancy of the extra factor $\Omega\bar{\Omega}^{-1}$ between (50) and (51) leads, via (28), to the presence of the torsion term in (49), we get, from (49),

$$\hat{\Psi}_{ABCD} = \Omega\bar{\Omega}^{-1}\Psi_{ABCD} \tag{53}$$

5. NON-LOCAL EFFECTIVE TORSION

Finally let us consider a certain physical argument which perhaps lends some support to the ECSK theory. The argument follows an earlier one that I gave some years ago[7] concerning a proposed relation between energy and focussing power in standard general relativity. That argument ran roughly as follows.

Imagine a small region $\mathscr{R}$ of Ricci curvature in an otherwise nearly flat Einsteinian (torsion-free) space-time $\mathscr{M}$. The rays (that is, idealized light rays, or null geodesics) which pass through $\mathscr{R}$ are affected in the same way that ordinary light would be as it passes through a positive lens, the focussing power of the lens being essentially proportional to the integrated energy tensor component which lies completely in the ray direction. If we, evisage two such regions $\mathscr{R}$ and $\mathscr{R}'$, and a ray γ which encounters both,

then we find that these energy contributions effectively add, as would powers of lenses, provided that the curvatures remain small, but that with larger curvatures there is a (negative) nonlinear correction term, as again there would be with ordinary lenses, which is proportional to the product of the two powers involved and the (affine) distance between the regions. This nonlinear contribution turns out to be consistent with what would be expected of a (negative) gravitational potential energy contribution to the total energy.

Next, consider a small region $\mathscr{W}$ of weak Weyl curvature. We now find that the effective lens is *purely astigmatic*—in the sense that the focussing in each direction orthogonal to the ray γ is the negative of the focussing in the perpendicular direction—whereas the effect of Ricci tensor was *anastigmatic*. If γ encounters two such weak Weyl regions $\mathscr{W}$ and $\mathscr{W}'$, then the resultant focussing is again (nearly) purely astigmatic, being the sum of the effects of each of $\mathscr{W}$ and $\mathscr{W}'$ separately. But for stronger curvature the nonlinear correction terms begin to play their part, and produce an effective additional positive anastigmatic contribution which appears, at large distances, like the effect of the presence of some Ricci tensor for a positive energy density. In particular a region of highly oscilliatory Weyl curvature in empty space ($R_{ab} = 0$) will produce a resultant effect which, on large "coarse-grained" scales, resembles the presence of Ricci tensor. The effective energy carried by the region of Weyl curvature, as measured using the Bondi–Sachs definition,[7,23] can be correctly computed at large distance using this focussing effect and equivalent effective Ricci tensor.

Next suppose that a ray γ encounters a region of Weyl curvature ($R_{ab} = 0$) for which the direction of astigmatism rotates rapidly along γ, rather than oscillates. We find, now, that in addition to an effective nonlocal positive focussing there is a resultant *rotational* effect. That is, there is an effective rotation of Lie transport along a small bundle of rays, as compared with parallel transport. This is an effect which cannot occur *locally* in (pseudo–) Riemannian space-time, but it resembles the local polarization-plane rotation effect induced by *torsion* that we discussed earlier.

We may imagine that the gravitational field might possess an effective (nonlocal) spin density, much in the way that it possesses an effective (nonlocal) energy density. I have indicated[7] how focussing power can indeed give a rough but plausible measure of this effective gratitational energy density in a way consistent with Einstein's theory. The question that now arises is whether this effective rotation along γ gives a plausible measure of the effective spin-density of gravitation in a way consistent with the ECSK theory.

To fix ideas, let us suppose that $\mathscr{M}$ is a plane-wave vacuum space-time ($R_{ab} = 0$, $T_{ab}{}^{c} = 0$) and that γ is a ray which cuts across the wave fronts.

Suppose also that γ belongs to a congruence of rays so that the complex "optical scalars" ρ, σ or Sachs may be defined.[28,7,23,8] Here ρ measures "convergence + $i \times$ rotation" of the congruence along γ, while σ measures its "shear" ($\frac{1}{2}\arg\sigma$ defining the direction of the shear axis). We define complex quantities ξ, η along γ, subject to

$$D\begin{pmatrix}\xi & \eta\\ \bar\eta & \bar\xi\end{pmatrix} = -\begin{pmatrix}\rho & \sigma\\ \bar\sigma & \bar\rho\end{pmatrix}\begin{pmatrix}\xi & \eta\\ \bar\eta & \bar\xi\end{pmatrix} \tag{54}$$

(with D defining propagation derivative along γ) which may be regarded as describing the "state" of the congruence in the neighborhood of γ, in direction orthogonal to γ. To see the significance of ξ and η, imagine an Argand plane Π, which is orthogonal to γ at each point of γ and parallelly propagated along γ. The rays neighboring to γ which are "abreast" with γ are those which meet one, and therefore all, of the Π's. As we move up γ these rays provide a continuous succession of linear mappings of Π to itself which can be described, if we start from some base point $P \in \gamma$, at which $\xi = 1$, $\eta = 0$, by

$$z \mapsto \xi z + \eta\bar z \tag{55}$$

i.e.,

$$\begin{pmatrix}z\\ \bar z\end{pmatrix} \mapsto \begin{pmatrix}\xi & \eta\\ \bar\eta & \bar\xi\end{pmatrix}\begin{pmatrix}z\\ \bar z\end{pmatrix} \tag{56}$$

Taking a standard null tetrad ℓ^a, m^a, $\bar m^a$, n^a, parallelly propagated along γ, with ℓ^a pointing along γ and with $Re(m^a)$ and—$Im(m^a)$, respectively, pointing along the real and imaginary axes of Π, we have the Sachs equations

$$D\begin{pmatrix}\rho & \sigma\\ \bar\sigma & \bar\rho\end{pmatrix} = \begin{pmatrix}\rho & \sigma\\ \bar\sigma & \bar\rho\end{pmatrix}^2 + \begin{pmatrix}\Phi & \Psi\\ \bar\Psi & \Phi\end{pmatrix} \tag{57}$$

and

$$D^2\begin{pmatrix}\xi & \eta\\ \bar\eta & \bar\xi\end{pmatrix} = -\begin{pmatrix}\Phi & \Psi\\ \bar\Psi & \Phi\end{pmatrix}\begin{pmatrix}\xi & \eta\\ \bar\eta & \bar\xi\end{pmatrix} \tag{58}$$

where

$$\Phi = \bar\Phi = -\tfrac{1}{2}R_{ab}\ell^a\ell^b = \Phi_{000'0'} \tag{59}$$

and

$$\Psi = R_{abcd}\ell^a m^b \ell^c m^d = \Psi_{0000} \tag{60}$$

825/13/3-4

For a circularly polarized gravitational plane-wave we can take

$$\Psi = Ae^{i\omega r}, \qquad \Phi = 0 \tag{61}$$

where r is an affine parameter along γ with

$$Dr = 1 \tag{62}$$

and $r = 0$ at P, and where A and ω are real constants describing the amplitude and $2\pi \times$ frequency of the wave, respectively, with $\omega > 0$ for a right-handed helicity. (This situation has been studied by Berk[29]). Let us suppose that the rays start off parallel at P, so $\xi = 1$ and $\eta = D\xi = D\eta = 0$ at $r = 0$. The resulting solution of (58), with (61), is

$$\xi = A\left\{\frac{e^{\frac{1}{2}i(\omega+p)r}}{p(\omega+p)} - \frac{e^{\frac{1}{2}i(\omega-p)r}}{p(\omega-p)} - \frac{e^{\frac{1}{2}i(\omega+q)r}}{q(\omega+q)} + \frac{e^{\frac{1}{2}i(\omega-q)r}}{q(\omega-q)}\right\} \tag{63}$$

with a similar expression for η, where

$$p = (\omega^2 + 4A)^{1/2}, \qquad q = (\omega^2 - 4A)^{1/2} \tag{64}$$

We assume that the rotation is of high frequency, with

$$\omega \gg 4A \tag{65}$$

The first and third terms in the curly bracket in (63) are swamped by the second and fourth. The behavior of ξ is initially to oscillate backwards and forwards along the real axis in the Argand plane with approximate frequency

$$A(2\pi\omega)^{-1} \tag{66}$$

(second and fourth terms) this motion being perturbed by small oscillations of frequency roughly

$$\omega(2\pi)^{-1} \tag{67}$$

(first and third terms). But gradually the direction of this main oscillation rotates ("beats" between the second and fourth terms) with frequency approximately

$$2A^2\pi^{-1}\omega^{-3} \tag{68}$$

(These frequencies are simply obtained by expanding $-\frac{1}{2}(\omega - p)$ or $\frac{1}{2}(\omega - q)$ to 1st order, $\frac{1}{2}(\omega + p)$ or $\frac{1}{2}(\omega + q)$ to 0th order, and $\frac{1}{4}(2\omega - p - q)$ to 3rd order in $A\omega^{-2}$, respectively.)

In thinking of the effect of these oscillations in the Argand plane Π, we shall ignore η, since that corresponds to the direct effect of astigmatism, which is to be averaged away in these considerations. The value of $|\xi|$ then provides the measure by which the rays have expanded and $\arg \xi$, the measure of their resultant rotation. A uniform presence of Ricci tensor with

$$\Psi = 0, \qquad \Phi = A^2 \omega^{-2} \tag{69}$$

would produce an oscillation like that of (66), so we conclude that (69) is a measure of the effective energy density in the wave. Similarly, a uniform torsion of strength proportional to

$$A^2 \omega^{-3} \tag{70}$$

would produce the rotation frequency given by (68).

We may compare these powers with what would be expected for the energy and angular momentum density for a massless spin-2 field whose complex amplitude is given by (61). For the energy density, we expect something of the general form

$$\left(\int \Psi \, dr\right)\left(\int \bar{\Psi} \, dr\right) = A^2 \omega^{-2} \tag{71}$$

and for the angular momentum density, something like

$$\frac{1}{2} i \left(\int\left(\int \Psi \, dr\right) dr\right)\left(\int \bar{\Psi} \, dr\right) - \frac{1}{2} i \left(\int \Psi \, dr\right)\left(\int\left(\int \bar{\Psi} \, dr\right) dr\right) = A^2 \omega^{-3} \tag{72}$$

This agreement between (71) and (69) was already clear from earlier work (cf. Penrose,[7] Berk[29]) and is part of the general consistency of Einstein's theory. The agreement between (72) and (70) seems to lend some support to the ECSK theory.

6. CONCLUSIONS?

The apparently satisfactory result for the ECSK theory that we have just obtained carries with it some problems, however. According to that theory, the spin density of a neutrino wave (spin-$\frac{1}{2}$) would manifest itself directly through torsion, and this would produce rotation in a parallelly propagated polarization plane along a ray. And as we have just seen, the effective (nonlocal) spin density of a gravitational wave (spin-2) similarly produces polarization plane rotation. (These effects are in principle

observable.[30]) But what about electromagnetism (spin-1)? Here there is no rotation effect from standard general relativity since by the duality rotation invariance of the Maxwell stress tensor, only the amplitude and not the frequency of a circularly polarized electromagnetic wave will show up in the curvature. But also there is no gauge invariant local expression for the spin density, so the normal procedure has been to take the torsion to vanish, and again we get no polarization plane rotation effect.

This discrepancy between spins-$\frac{1}{2}$ and 2, on the one hand, and 1, on the other, seems puzzling. Perhaps there is some way of coupling together the two main ideas described in this article. For example, it might be that certain kinds of torsion are not locally observable, and then one would not mind if an electromagnetic gauge transformation affected only this unobservable part of the torsion. It is tempting to suggest, say, that the change in torsion which is produced by a complex conformal transformation might for some reason be unobservable. Unfortunately this does not seem to help with the problem with electromagnetism and we are left with something of an enigma. Some more subtle interrelations appear to be required.

REFERENCES

1. P. A. M. Dirac, *Proc. Roy. Soc.* **A117**, 610–624 (1928); **A118**, 351–361 (1928).
2. E. Cartan, *Ann. Ec. Norm. Sup.* **40**, 325–412 (1923); **41**, 1–25 (1924); **42**, 17–88 (1925).
3. D. W. Sciama, in *Recent Developments in General Relativity* (Pergamon and PWN, Oxford, 1962), pp. 415–439.
4. T. W. B. Kibble, *J. Math. Phys.* **2**, 212–221 (1961).
5. A. Trautman, *Bull. Acad. Pol. Sci.* **20**, 185–190, 503–506, 895–896 (1972); *Ser. Sci. Math. Astron. Phys.* **21**, 345–346 (1973).
6. F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester, *Rev. Mod. Phys.* **48**, 393–416 (1976).
7. R. Penrose, in *Perspectives in Geometry and Relativity*, B. Hoffman, ed. (Indiana University Press, Bloomington, 1966), pp. 291–309.
8. R. Penrose and W. Rindler, *Spinors and Space-Time*, Vol. I (Cambridge University Press, Cambridge, 1983).
9. L. Infeld and B. L. van der Waerden, *Sitzungsber. Preuss. Akad. Wiss. Physik.-math.*, **K1**(9), 380–401 (1933).
10. P. A. M. Dirac, *Proc. Roy. Soc.* **A115**, 447 (1936).
11. M. Fierz, *Helv. Phys. Acta* **13**, 45–60 (1940).
12. H. Weyl, *Z. Phys.* **56**, 330–352 (1929).
13. O. Laporte and G. E. Uhlenbeck, *Phys. Rev.* **37**, 1380 (1931).
14. H. A. Buchdahl, *Nuovo Cim.* **10**, 96–103 (1958).
15. J. Plebanski, *Acta Phys. Polon.* **27**, 361–393 (1965).
16. R. Penrose, *Ann. Phys.* (New York) **10**, 171–201 (1960).
17. S. Deser and B. Zumino, *Phys. Lett.* **62B**, 335–337 (1976).
18. E. Cunningham, *Proc. Lond. Math. Soc.* **8**, 77 (1910).
19. H. Bateman, *Proc. Lond. Math. Soc.* **8**, 223 (1910).

20. P. A. M. Dirac, *Ann. Math.* **37**, 429–442 (1936).
21. J. A. McLennan, Jr., *Nuovo. Cim.* **10**, 1360 (1956).
22. R. Penrose, *Proc. Roy. Soc. London* **A284**, 159–203 (1965).
23. R. Penrose, in *Battelle Recontres, 1967 Lectures in Mathematics and Physics*, C. M. DeWitt and J. A. Wheeler, eds. (Benjamin, New York, 1968).
24. H. A. Buchdahl, *Nuovo Cim.* **11**, 496–506 (1959).
25. R. Penrose and R. S. Ward, in *General Relativity and Gravitation, One Hundred Years after the Birth of Albert Einstein*, Vol. II, A. Held, ed. (Plenum Press, New York, 1980).
26. M. G. Eastwood, R. Penrose, and R. O. Wells, Jr., *Commun. Math. Phys.* **78**, 305–351 (1981).
27. W. T. Payne, *Am. J. Phys.* **20**, 253–262 (1952).
28. R. Sachs, *Proc. Roy. Soc.* **A264**, 309 (1961).
29. A. A. Berk, *Gravitational (and other) Lens Systems* (Ph.D. Dissertation, Birkbeck College, University of London, 1978).
30. R. Matzner and B. W. Tolman, *Phys. Rev.* **D26**, 2951 (1982).

141
NON-LOCALITY IN PHYSICS: A NEW MATHEMATICAL VIEWPOINT

International workshop on present thinking about parts and wholes. Lund, Sweden, 1–3 June 1983

Non-locality in physics:
a new mathematical viewpoint

by

Roger Penrose
Mathematical Institute, Oxford

(Paper written for the International Workshop:
"Present Thinking about Parts and Wholes";
Lund, Sweden, 1 - 3 June, 1983.)

Perhaps the most profound and influential innovation in the dynamics of Galilei-Newton was that the motions of bodies are best described in terms of the differential equations which govern their behaviour from moment to moment, rather than in terms of the entire geometric configuration traced out by their motions as a whole. The true simplicity lying behind the complicated detailed motions of planets and projectiles could be revealed only by examining their (infinitely) small-scale behaviour. Later on, Maxwell's electromagnetic theory forced physicists to appreciate the fundamental role played by physical fields, whose behaviour is governed by partial differential equations, rather than by the ordinary differential equations of the older Newtonian theory. Again, it was only by studying the (remarkably elegant) mathematical behaviour

of these fields on the (infinitely) small scale that the essential unity of some seemingly diverse phenomena (this time magnetism, electricity and light) could be understood.

Partial differential equations are also fundamental ingredients of the two most profound upheavals in physical understanding that have taken place in this century: Einstein's general relativity, and the quantum theory. So again the basic equations of these theories are concerned with behaviour on an (infinitely) small scale. But now some non-local ingredients have begun to creep in. In general relativity, statements must be made about the global structure of solutions of Einstein's equations in order that they be physically acceptable (such as the non-existence of closed timelike curves, so that the worst type of causality violation may be avoided). The differential equations by themselves are no longer quite sufficient to characterize a physical model completely. Moreover, certain important physical quantities acquire non-local ingredients in Einstein's theory, the most noteworthly of these being energy, the gravitational contributions to which cannot be localized. In quantum mechanics, non-locality plays an even more essential role. This shows up in quantum correlations between widely separated particles, and in the effective "instantaneous state-vector reduction" which is part of the standard interpretation of the theory with regard to the observation process. Such non-locality shows up in a striking way in the Einstein-Podolsky-Rosen

thought experiment and in the observed violation of the inequality which Bell showed would have to hold in any local (classical-type) theory.

That there are basic non-local ingredients to physical phenomena seems now to be experimentally well substantiated. Such ingredients present puzzling aspects to the standard space-time picture of the world. Partly as an attempt to resolve such puzzles I have, for a number of years, been attempting to promote an alternative geometric picture for the description of physical phenomena. In this picture, the standard four-real-dimensional space-time geometry of relativity is transcribed into a certain quite different-looking complex geometry (i.e. a geometry based on complex numbers: $a + b\sqrt{-1}$) by means of a non-local correspondence. This is the geometry of twistor theory.

In the twistor picture, the basic ingredient of physical geometry is not a space-time point (i.e. "event"), but a mathematical construct (a twistor) which may be viewed in space-time terms as the entire history a spinning object which travels in one direction with the speed of light. It is neither localized in space nor in time, though it represents a single point in twistor space. Conversely, a single space-time point may also be represented in twistor terms, not as a localized object, but as a two-complex-dimensional (linear) subspace of twistor space.

Linear physical fields (at least massless ones, such as Maxwell's electromagnetic field) turn out to be neatly describable in twistor terms, by means of a remarkable

family of mathematical concepts referred to as sheaf cohomology theory. This theory was developed originally (by Leray) for quite a different purpose, and it later came into its own as an essential ingredient of the modern theory of several complex variables. In a certain sense it may be regarded as providing a holistic calculus. The local mathematical concept of "function" is generalized to yield non-local higher-order kinds of object (elements of sheaf cohomology groups) which, when localized to too small a region may vanish completely, and which describe more subtle kinds of non-local structure. It seems that such objects may provide an important ingredient for the physical descriptions of a future theory and are, in any case, well worth studying for their own sake.

Moreover, sheaf cohomology is related in an intimate way to certain non-linear mathematical constructs, notably fibre bundle theory and the global construction of manifolds, which play significant roles in modern physical theories (in Yang-Mills theories of particle interactions and in general relativity and its modifications). In twistor theory this intimate relationship has shown up in a remarkable (but as yet incomplete) way. The local space-time values of a physical field (Yang-Mills field or space-time curvature) are codified in the global structure of the corresponding twistor space. And the partial differential equations that these field quantities are required to obey turn out to be automatically satisfied by virtue of this fact! As yet, this procedure works only for certain classes

of such fields (self-dual or stationary), but there is considerable hope that it will eventually be possible to handle the general case also in this way. It is, indeed, one of the more ambitious aims of the whole twistor programme to replace all the basic differential equations of physics by statements about global structure. If such aims could eventually be realized, then a fundamentally holistic ingredient would be seen to be present in the determination of seemingly local physical laws.

142
ON CLOSED-SET COVERINGS FOR PT^+—A CORRECTION

Twistor Newsletter **16** (1983). Also published in *Further Advances in Twistor Theory,* Vol. 1. Eds L. J. Mason and L. P. Hughston. Longman, 1990

This corrects the error in Chapter 81, found also in the final part of Chapter 77.

On Closed-set coverings for PT⁺—A correction. *Twistor Newsletter* **16** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

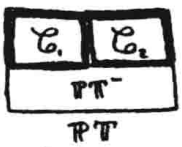
On Closed-set Coverings for $\overline{\mathbb{PT}^+}$ — A Correction

In a note in TN 3, p. -1 (unfortunately also reprinted in Adv. in Twistor Thory, pp. 35, 36) I made the suggestion that the closed-set covering of $\overline{\mathbb{PT}^+}$ given by the 2 closed-sets

$$\mathcal{C}_1 : |Z^0 + Z^2| \geq |Z^1 + Z^3|, \; Z^\alpha \bar{Z}_\alpha \geq 0$$

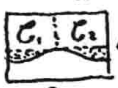
and

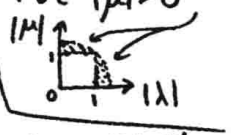
$$\mathcal{C}_2 : |Z^0 + Z^2| \leq |Z^1 + Z^3|, \; Z^\alpha \bar{Z}_\alpha \geq 0$$



would suffice for all positive-frequency analytic massless wave functions (defined on $\overline{\mathbb{CM}^+}$), claiming that an explicit inverse twistor function construction would show this. However this is actually not so, it being necessary that the field be defined in a somewhat larger region, and the claim, as stated, is false

Any twistor function defined in a neighbourhood of $\mathcal{C}_1 \cap \mathcal{C}_2$ would yield a field at any point of $\mathbb{CM}$ corresponding to a line in $\mathbb{PT}$ which meets $\mathcal{C}_1 \cap \mathcal{C}_2$ in an S^1. It is not hard to see that, for example, the line joining $(1+\lambda, 0, 1-\lambda, 0)$ to $(0, 1+\mu, 0, 1-\mu)$ has this property provided that $|\lambda^2| + |\mu^2| \leq 2$, whereas it extends into $\mathbb{PT}^-$ whenever $|\lambda| > 1$ or $|\mu| > 0$. So the resulting field extends to a fixed region outside $\mathbb{CM}^+$.





But it is clear that not all massless wave functions defined on $\mathbb{CM}^+$ can extend to such a region (e.g. the positive frequency part of a nowhere analytic field on the real compactified space $\mathbb{M}$ does not extend outside $\mathbb{CM}^+$, and by an arbitrarily small conformal transformation, $\mathbb{CM}^+$ can be transformed to $\overline{\mathbb{CM}^+}$ together with an arbitrarily narrow open "collar" extending into $\mathbb{CM}^-$). No such simple 2-set cover will work generally (and to have a reasonable chance might have to be based on such as the "snail contour" of TN 1., p. 22, BDB & RP. — unfortunately not reprinted in A.T.T.).

Thanks to Mike Eastwood for his perceptive scepticism and to Alex Pilato for reminding me of this problem. Roger Penrose

143
REMARKS ON CURVED-SPACE, TWISTOR THEORY AND GOOGLIES

Twistor Newsletter **16** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 Remarks on curved-space, twistor theory and googlies. *Twistor Newsletter* **16** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.11 **Remarks on Curved-Space Twistor Theory and Googlies** *by R. Penrose* (TN 16, August 1983)

Let $\mathscr{M}$ be a complex manifold (though much of what follows works in the real case also). Let $T^*(\mathscr{M})$ be its cotangent bundle, with naturally defined 1-form

$$\Phi = p_a dx^a$$

(cf. §III.4.10) with regard to notation) and a symplectic 2-form

$$\Theta = d\Phi = dp_a \wedge dx^a.$$

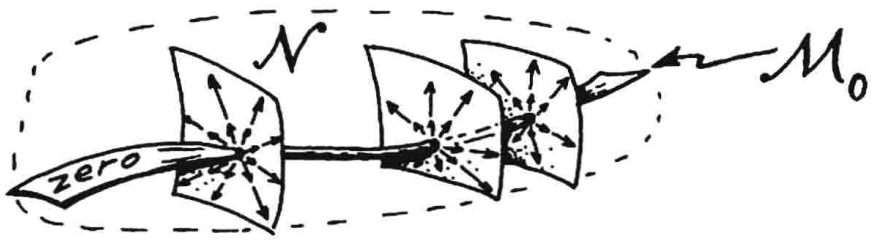
Let $\mathscr{N}$ be the manifold $T^*(\mathscr{M})$ but with the zero-section $\mathscr{M}_0$ ($\simeq \mathscr{M}$) removed, the structure of $\mathscr{N}$ being given by Φ. The Poisson structure (dual to Θ)

$$\{\ ,\ \} = \frac{\partial}{\partial p_a} \wedge \nabla_a,$$

together with Φ serves to define the vector field

$$\Upsilon = p_a \frac{\partial}{\partial p_a}.$$

If $\mathscr{M}_0$ is re-adjoined to $\mathscr{N}$, we can use the integral curves of Υ, 'tied together' at the points of $\mathscr{M}_0$, to reconstruct the fibres of $T^*(\mathscr{M})$, and hence to find the bundle structure of $\mathscr{N}$ over $\mathscr{M}$. But without $\mathscr{M}_0$, the local structure of $\mathscr{N}$ is insufficient.



A striking example occurs with projective flat twistor space $\mathscr{M} = \mathbb{PT}$. The cotangent bundle $T^*(\mathbb{PT})$ can be described as the space of (Z^α, W_α) under the equivalence

$$(Z^\alpha, \lambda W_\alpha) \sim (\lambda Z^\alpha, W_\alpha), \text{ where } Z^\alpha W_\alpha = 0, \ Z^\alpha \neq 0,$$

(so $\mathscr{N}$ is also a 1-bundle over ambitwistor space) and we have

$$\Phi = W_\alpha dZ^\alpha.$$

(This needs to be invariant under $Z^\alpha \mapsto \lambda(Z^\beta)Z^\alpha$, whence the conditions above.) Note that when we remove $\mathscr{M}_0$, so $W_\alpha \neq 0$ in addition to $Z^\alpha \neq 0$, then we get a space $\mathscr{N}$ which is completely *symmetrical* between $\mathbb{PT}$ and $\mathbb{PT}^*$. Were we to adjoin $\mathbb{PT}^*$, as $\mathscr{M}_0$, rather than $\mathbb{PT}$, then we should obtain the cotangent bundle of $\mathbb{PT}^*$ rather than of $\mathbb{PT}$. (Note that $\Phi = W_\alpha dZ^\alpha = -Z^\alpha dW_\alpha$ and $\Upsilon = Z^\alpha \frac{\partial}{\partial Z^\alpha} = W_\alpha \frac{\partial}{\partial W_\alpha}$). This is despite the fact that the adjoined zero-section fits in exactly the same 'hole' in $\mathscr{N}$ in each case! (An *essentially non-Hausdorff* space would arise if *both* zero-sections were adjoined at once.)

Similar situations arise with any (complex) space-time 4-manifold $\mathscr{Q}$, where now $\mathscr{N}$ is to be the space of its null geodesics, with scalings. The 'scalings' here mean that a (null) covector p_a is to be associated with the null geodesic, where p_a points along the geodesic and is parallely propagated along it (a conformally invariant property). The operator Υ generates changes in this scaling:

$$\Upsilon = p_a \frac{\partial}{\partial p_a},$$

Remarks on curved-space, twistor theory and googlies. *Twistor Newsletter* **16** (1983).
Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

and we have $\Phi = p_a dx^a$. Working locally (in the sense of near one of the integral curves of Υ) we find that many varieties of different zero-section can be adjoined to $\mathscr{N}$. Most notably, if we select a hypersurface $\mathscr{S}$ in $\mathscr{Q}$, then $\mathscr{N}$ may be regarded as effectively the cotangent bundle of $\mathscr{Q}$ (restrict p_a to tangent directions in $\mathscr{Q}$), or of the projective hypersurface twistor space $\mathbb{P}\mathscr{T}(\mathscr{S})$, or of the projective hypersurface dual twistor space $\mathbb{P}\mathscr{T}^*(\mathscr{S})$, the only distinction, in each case, being the different choices of adjoined zero-section. We can move $\mathscr{S}$ to $\mathbb{C}\mathscr{I}^+$ and regard $\mathscr{N}$ as the cotangent bundle of $\mathbb{C}\mathscr{I}^+$ or of projective asymptotic twistor space $\mathbb{P}\mathscr{T}$ or projective asymptotic dual twistor space $\mathbb{P}\mathscr{T}^*$. All the structural information seems to lie in the way that the zero-section sits in the 'hole' in $\mathscr{N}$.

This symmetry between $\mathbb{P}\mathscr{T}$ and $\mathbb{P}\mathscr{T}^*$ suggests that the above viewpoint may be helpful for the googly problem. Indeed, the following procedure seems quite hopeful. Consider, first the case of $\mathbb{T}$. Ignoring a sign ambiguity we can embed $\mathbb{T}^I$ (that is, $\mathbb{T}$ but with the region $\pi_{A'} = 0$ removed) in $T^*(\mathbb{P}\mathbb{T})$ by

$$\dot{Z} \mapsto (\dot{Z}, Z\rceil)\,.$$

Note that the *scale* of $\dot{Z}$ (up to a sign) is encoded in the scaling of the covector $W = Z\rceil$ (with $(\dot{Z}\dot{W}) \equiv (\lambda\dot{Z}, \lambda^{-1}\dot{W})$. Effectively equivalent is the assignment

$$\dot{Z} \mapsto \dot{Z}Z\rceil.$$

In place of I, we have a *blown-up* projective space (with I replaced by a quadric in $\mathbb{P}\mathbb{T}$), since if

$$\dot{Z} = \alpha^{-1}\omega\rfloor + \alpha\dot{o} \qquad (\omega, \dot{o} \text{ fixed}),$$

then

$$\dot{Z}Z\rceil = \frac{\omega\rfloor}{o\rceil} + \alpha^2\dot{o}\,o\rceil \longrightarrow \frac{\omega\rfloor}{o\rceil}$$

as $\alpha \to 0$, so both a point on I ($\omega\rfloor$) and a plane through it ($o\rceil$) are retained in the limit. Because of singular behaviour of googly maps at I, in the curved case, we seem to require a blow-up. The idea, here, is to adjoin 3-manifold $\mathbb{I}$ to $\mathbb{T}$, which consists just of these above limits.

Circumstantial evidence that finding deformations of the way that $\mathbb{I}$ is adjoined to $\mathbb{T}$ may be the right thing to do is provided by the following facts: Whereas the natural 4-form

$$d^4Z = \frac{1}{24}\,dZ \wedge dZ \wedge dZ \wedge dZ \quad \text{and 3-form} \quad Z \wedge dZ \wedge dZ \wedge dZ$$

have *singular* limiting values at $\mathbb{I}$, the forms

$$Z\rfloor Z\rfloor d^4Z \quad \text{and} \quad Z\rfloor Z\rfloor \mathscr{D}Z$$

are well-defined (and non-zero) there. (To see this it seems simplest to choose local coordinates at $\mathbb{I}$, e.g. $X = \omega^0\pi_{0'}$, $Y = \omega^0\pi_{1'}$, $Z = \omega^1\pi_{0'}$, $W = \pi_{0'}\pi_{0'}$.) Note that

$$d\{f(\dot{Z})Z\rfloor Z\rfloor \mathscr{D}Z\} = ((\Upsilon + 6)f)Z\rfloor Z\rfloor d^4Z)$$

so that if these forms are to be deformed (infinitesimally?) at $\mathbb{I}$ by multiplication by some twistor function, it would seem that those of homogeneity -6 have a special role to play. (Compare the 'googly photon' construction of Penrose §O.3.7 and Hughston §O.3.8.) In the curved case, each '$\dot{Z}Z\rfloor$' corresponds to a scaled null geodesic on $\mathbb{C}\mathscr{I}^+$ (lying in a β-plane on $\mathbb{C}\mathscr{I}^+$). The points of $\mathbb{I}$ correspond to scaled generators of $\mathbb{C}\mathscr{I}^+$.

It would seem that the role of the 'zeros' in relation to $\mathbb{I}$ must be important.

144
A NOTE ON SPARLING'S 3-FORM

Twistor Newsletter **16** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 A note on Sparling's 3-form. *Twistor Newsletter* **16** (1983). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.10 **A note on Sparling's 3-form** *by R. Penrose* (TN 16, August 1983)

This is just to draw attention to the remarkable fact noticed by George Sparling, §III.3.8 & §III.3.7, that the 3-form

$$\Gamma = id\pi_{A'} \wedge d\bar{\pi}_A \wedge dx^{AA'}$$

defined on the spin-vector bundle $(x; \pi_{A'}, \bar{\pi}_A)$ of a curved space-time is *closed* iff $\mathscr{M}$ *is an Einstein vacuum*, in which case it is the exterior derivative of the 2-form

$$\Omega = i\pi_{A'} d\bar{\pi}_A \wedge dx^{AA'}$$

(or of the Witten-Nestor 2-form $\frac{1}{2}(\Omega + \bar{\Omega})$. (Sparling would write θ^a for my dx^a and $\theta_{A'}$ for my $d\pi_{A'}$, etc. Here $d\pi_{A'}$ is indeed the "exterior derivative" of $\pi_{A'}$, though $d^2 \neq 0$ on $\pi_{A'}$ since $\pi_{A'}$ has an abstract index. There is no such thing as "x^{a}" with an abstract index, but $dx^a = \delta^a_{\mathbf{a}} dx^{\mathbf{a}}$. We have $d^2x^a = 0$ if the torsion vanishes; otherwise $d^2x^a = -\frac{1}{2}T^a_{bc}dx^b \wedge dx^c$.

Note that this seems to fit in with a certain pattern of significance in twistor theory: $\Sigma = Z^\alpha \bar{Z}_\alpha$ and $d\Sigma = i\bar{\Phi} - i\Phi$ where $\Phi := iZ^\alpha d\bar{Z}_\alpha = p_a dx^a = \pi_{A'}\bar{\pi}_A dx^a$ on $\Sigma = 0$. Then $\Theta := d\Phi = idZ^\alpha \wedge d\bar{Z}_\alpha = dp_a \wedge dx^a = d(\pi_{A'}\bar{\pi}_A) \wedge dx^a = i\bar{\Omega} - i\Omega$ and $d\Omega = \Gamma$ in vacuum.

145
PHYSICAL SPACE-TIME AND NONREALIZABLE CR-STRUCTURES

Bull. Amer. Math. Soc. (N.S.) **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

Twistor theory has provided ways in which sophisticated mathematics can sometimes give new insights into physical behaviour. This article (presented in honour of Henri Poincaré) shows how the debt can be partly repaid, whereby physical understandings can provide new insights into mathematical problems. Specifically, the theory shows us how to construct a complex manifold with a C^∞-smooth boundary that cannot be extended as a complex manifold beyond that boundary.

282 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society,* Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

PHYSICAL SPACE-TIME AND NONREALIZABLE CR-STRUCTURES

BY ROGER PENROSE

Abstract. Space-time views leading up to Einstein's general relativity are described in relation to some of Poincaré's early ideas on the subject. The basic geometry of twistor theory is introduced as it arises both from Minkowski space-time and the more general curved Einstein models. It is shown how this provides a CR-structure (this being, in essence, another of Poincaré's pioneering concepts) in a natural way. Nonrealizable CR-structures can arise, and an example is presented, due to C. D. Hill, G. A. J. Sparling and the author, of a complex manifold-with-boundary which cannot be extended as a complex manifold beyond its C^∞ boundary.

1. Introductory remarks. The nature of physical geometry was something that held considerable interest for Poincaré, and he often referred to it in his more popular writings. Moreover, it was Poincaré who first clearly understood the physical transformation group of special relativity—arising as a group of symmetries of Maxwell theory—and he suggested that the relativity principle concerned might hold for physics generally. This dates back to 1899, six years before Einstein's first paper on relativity (cf. Poincaré (1906), (1954) for details). It is fitting, therefore, that this symmetry group should be now very commonly referred to as the Poincaré group (otherwise known as the inhomogeneous Lorentz group, where simply "Lorentz group" now normally refers only to the related homogeneous group).

Poincaré also had interesting and, to some extent, insightful things to say about the possibility that physical space might have a non-Euclidean geometry. But here his instincts appear ultimately to have let him down. For in "Science and Hypothesis" Poincaré (1905) wrote

> "What then, are we to think of the question: Is Euclidean geometry true? It has no meaning. We might as well ask if the

---

This lecture was delivered at the Poincaré Symposium, April 1980; received by the editors June 1, 1982.

1980 *Mathematics Subject Classification.* Primary 83C05, 53C20.

> metric system is true, and if the old weights and measures are false; if Cartesian coordinates are true and polar coordinates false. One geometry cannot be more true than another; it can only be more convenient. Now, Euclidean geometry is, and will remain, the most convenient: 1st, because it is the simplest, and it is not so only because of our mental habits or because of the kind of direct intuition that we have of Euclidean space; it is the simplest in itself, just as a polynomial of the first degree is simpler than a polynomial of the second degree; 2nd, because it sufficiently agrees with the properties of natural solids, those bodies which we can compare and measure by means of our senses."

It is, of course, no discredit to Poincaré that he should have had no inkling of that extraordinary development, which emerged three years after his death, namely Einstein's *general* theory of relativity. Nevertheless, I find it rather remarkable that Poincaré, with his deep geometric and philosophical insights, should apparently have had such a rigid feeling that geometries other than Euclidean could have no chance of providing accurate and useful descriptions of physical space according to some future theory. His remarks on the conventionality of geometry and on the fact that the geometry we use in physics is of necessity an idealization are, indeed, quite profound and insightful. But he seems to have made a rather bad mistake in his reasoning, his critical faculties being perhaps dulled by an erroneous (but natural enough) intuitive presumption that in "our world" the geometry actually *is* Euclidean!

According to modern cosmology, it is quite on the cards that the large-scale spatial geometry of the universe may indeed be closely in accord with Lobachevskian geometry—that geometry which held such a fascination for Poincaré the mathematician, yet which had been rejected as inevitably physically inappropriate by Poincaré the philosopher! It would be interesting to know how Poincaré would have reacted to present-day cosmology. Perhaps there is an object lesson here for all of us.

Even if observation finally turns against its present slight preference for the Lobacheveskian spatial geometry—and also tells against models with a positively curved spatial geometry—leading us to believe that the large-scale spatial structure is Euclidean after all, we cannot now go back to flatness for the structure on a more localized level. Tests of general relativity are now sufficiently good to provide direct measurements of deviations from flatness in physical geometry, the geometry being defined in a sense that I shall describe in the next section. However, it is the geometry of *space-time*, rather than that of space, which has a clear physical interpretation. And we shall see that the space-time geometry of (Poincaré's!) *special* relativity, though flat, is yet *not* the geometry of Euclid.

284 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

In §2, I shall outline the contemporary and now well-established (Einstein) view of curved space-time geometry. Then in §3, I shall indicate how this geometry may be looked at in a different way, and how this new viewpoint enables another of Poincaré's pioneering innovations, namely the study of the instrinsic structure of boundaries of complex domains, to be applied in an unexpected context in physics. Finally, in §4, I shall indicate how, in a sense, these ideas also enable the physics to repay its debt to the mathematics and provide what appears to be the first established example of a complex manifold which cannot be locally extended beyond its C^∞-smooth boundary.

2. Structure of space-time. It is to Minkowski (1908) whom we owe the idea that physical geometry should be a 4-dimensional space-time geometry, rather than a 3-dimensional spatial one. In relativity theory, there is no absolute 3-geometry, but the 4-geometry of space-time is an absolute objective physical structure. In fact, when viewed in the Minkowskian way, relativity theory becomes a theory of the absolute. What have become relative are merely the older ideas of separate space and time. One further point about Minkowski's absolute space-time geometry that will emerge is that it is most directly described in terms of measured time-intervals rather than of distances in the ordinary sense, so that the geometry is really a "chronometry". This point of view has been emphasized particularly by Synge (1960) and Bondi (1965) and it leads to a considerable clarification of the meanings of the mathematical structures involved.

In order to appreciate fully the nature of the space-time provided by Einstein's general relativity, it may be helpful first to see how the older physical theories can be described in a space-time setting. We shall consider, in turn, five alternative space-time structures, which I shall refer to (cf. Penrose (1968)) as

A: Aristotelean space-time,

G: Galilean space-time,

N: Newtonian space-time,

M: Minkowskian space-time,

E: Einsteinian space-time.

The structure of the Aristotelean space-time A is given by its expression as a product

$$A = \mathbb{E}^1 \times \mathbb{E}^3, \tag{2.1}$$

where $\mathbb{E}^n$ denotes Euclidean n-space equipped with its flat metric (and, if desired, with an orientation), whence each $\mathbb{E}^n$ possesses a $\frac{1}{2}n(n+1)$ group of symmetries. For any two events $a, b \in A$, there is both an absolute time-difference $T(a, b) \in \mathbb{R}$ and an absolute spatial separation $D(a, b) \in \mathbb{R}$ (these being the distance functions in $\mathbb{E}^1$ and $\mathbb{E}^3$, respectively). It is therefore meaningful to say, of two events, whether they are simultaneous in A (i.e. $T(a, b) = 0$); but it is also meaningful to say whether they have the same position (i.e. $D(a, b) = 0$) irrespective of their simultaneity. For in Aristotelean physics the state of rest is distinguished among all motions, and this is reflected in the fact that the canonical copies of $\mathbb{E}^1$ in

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

$\mathbb{E}^1 \times \mathbb{E}^3$ constitute a distinguished family of curves in A representing the world-lines of particles at rest, i.e. which have the same position at all times. The space A has a 7-parameter symmetry group.

The structure of Galilean space-time G, on the other hand is not that of a product but merely a fibre bundle with $\mathbb{E}^3$ fibres and an $\mathbb{E}^1$ base space, the projection map

$$G \to \mathbb{E}^1 \tag{2.2}$$

being the assignment of a "time" to any event in G. Thus, of any two events $a, b \in G$, it is meaningful to speak of their time-difference $T(a, b) \in \mathbb{R}$ and to say when the events are simultaneous (i.e. $T(a, b) = 0$), but their spatial separation *only* has meaning if they *are* simultaneous. This reflects the fact that in Galilean physics there is no invariantly defined state of rest, and any two nonsimultaneous events will be viewed as having the same spatial location with respect to *some* suitably moving reference system. The bundle structure of G is not, however, G's whole structure. We need also to single out a family of curves in G which represent *inertially moving particles*; and these curves may be referred to as *straight lines* in G. The simplest way to specify the particular straight line structure that G possesses is, perhaps, to say that G is an *affine space* and that the fibration (2.2) is *compatible* with this affine structure (in the sense that the fibres are all parallel, that the $\mathbb{E}^3$ fibres and $\mathbb{E}^1$ base all inherit their correct affine structures, and that parallel straight lines, transverse to the $\mathbb{E}^3$'s, provide metric-preserving maps between the $\mathbb{E}^3$'s). The total structure that all this assigns to G is somewhat weaker than the structure of A. The symmetry group of G contains that of A as a subgroup, and it is a 10-parameter group referred to as the *Galilei group*.

We pass now to Newtonian space-time N, the idea here (due to É. Cartan (1923), (1924), cf. also Friedrichs (1928), Trautman (1966)) being to treat Newtonian gravitational theory as a geometric theory in the spirit of Einstein's general relativity. We recall, for example, that a *uniform* Newtonian gravitational field permeating the whole of space would be totally undetectable. All particles would accelerate together under this field, so that relative to the particles themselves, the field would appear not to be there at all. It is only deviations from uniformity, i.e. "tidal forces", which are physically detectable, and the idea is that in the general case these supply a kind of curvature for N. As a fibre bundle, N has a structure identical with that for G,

$$N \to \mathbb{E}^1 \tag{2.3}$$

(so simultaneity and time-difference, and lack of a spatial separation concept for nonsimultaneous events, are the same as before)—but now there is a different family of curves representing inertial motions. "Inertial motion" now means "free motion under gravity", and we can refer to the corresponding world-lines in N as "geodesics". In fact, while not extremal curves, these geodesics are, indeed, the geodesics of a certain torsion-free linear connection Γ on N. The space N also

possesses a *co*metric $\mathbf{g}^*$ (symmetric contravariant 2-valent tensor) which is semi-definite but degenerate (having matrix-rank 3), where $\mathbf{g}^*$ has the defining property of inducing the correct Euclidean metric on each $\mathbb{E}^3$ fibre. The connection Γ preserves $\mathbf{g}^*$, but this property does not define Γ uniquely. The different possible inequivalent Γ's correspond to the different possible physically inequivalent gravitational fields. Thus, unlike G and A, which have canonical structures, there are many *different* Newtonian space-times N. Indeed, G is a special case, given when Γ is flat. In general, N has no symmetries.

Minkowskian space-time M, the space-time of special relativity, on the other hand, possesses, like G and A, a canonical and uniform structure. Like G (and like A), M is an affine space, and inertial motions in M are described by lines which are straight with respect to this affine structure. These lines may also be thought of as the geodesics of a certain flat connection Γ, where Γ is now the unique torsion-free connection preserving a certain flat nondegenerate pseudometric $\mathbf{g}$, of Lorentzian signature $(+,-,-,-)$. (It is the nondegeneracy of $\mathbf{g}$ that now enables it to determine Γ uniquely.) The pseudometric $\mathbf{g}$ serves to define a quadratic squared "distance" function S of signature $(+,-,-,-)$, on pairs $a, b \in M$ ($S(a, b) \in \mathbb{R}$). Thus M is a pseudo-Euclidean space. The physical meaning of S is that an (ideal) inertial (i.e. unaccelerated) clock which moves from the event a to the event b will register a time-interval between a and b equal to $\{S(a, b)\}^{1/2}$. For a physical clock (or, indeed, for any massive particle) $S(a, b) > 0$, and we say that the separation between a and b is *timelike*. When $S(a, b) = 0$ we say that the separation is *null*, this being the case when a light ray connects a to b. When $S(a, b) < 0$ we say that the separation is *spacelike*, and in this case a and b will appear as simultaneous in some suitably moving inertial reference system, the distance between them (using units for which the light-speed is unity) being $\{-S(a, b)\}^{1/2}$. The symmetry group of M is the 10-parameter *Poincaré group*.

The structure of Einsteinian space-time E bears the same relation to M as N does to G. Thus there are many inequivalent spaces E, their different structures providing the various inequivalent gravitational fields. The structure, in each case, is given by a pseudometric $\mathbf{g}$, with the same $(+,-,-,-)$ signature as M, but now $\mathbf{g}$ is generally not flat. The inertial motions in E (i.e. free motions under gravity) are given, as with M, by geodesics of the unique torsion-free connection Γ preserving $\mathbf{g}$. But as with N, a *curvature* of Γ provides the physically detectable "tidal force". As with M, the pseudometric $\mathbf{g}$ provides the definition of time-interval as measured by ideal clocks, but because $\mathbf{g}$ is not flat, we normally think of the interval as defined only between infinitesimally separated points, the interval (in the timelike or null case) being given by $\{\mathbf{g}(\delta x, \delta x)\}^{1/2}$, where δx is the tangent vector which "connects" the point to its infinitestimally separated neighbour. The world-line of a massive particle has tangent vectors which are all timelike (i.e. $\mathbf{g}(\delta x, \delta x) > 0$), the tangent vectors being all null (i.e. $\mathbf{g}(\delta x, \delta x) = 0$) in the case of a massless particle. The time-interval between finitely-separated events a and b, as

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

measured by a clock carried from a to b by such a particle is the pseudo-Riemannian "distance", given by

$$\int_a^b \{\mathbf{g}(dx, dx)\}^{1/2}.$$

Note that this time-interval depends upon the path through space-time from a to b. If a and b are not too far apart, this time-interval is a *maximum* for the geodesic (i.e. inertial path) which locally connects a to b. The metric $\mathbf{g}$ determines (via Γ) a Riemann curvature tensor $\mathbf{R}$. The direct physical interpretation of $\mathbf{R}$ is obtained from the Jacobi equation, which describes, in terms of $\mathbf{R}$, the geodesic derivation of (say) timelike geodesics, i.e. the tidal effect on inertially moving particles. In the general case, E has no symmetries.

In order to provide a physical theory of gravity, Einstein had to supplement the above general geometric structure by his *field equations*. These state that the trace-reversed Ricci tensor constructed from $\mathbf{R}$ is a constant multiple of the physical energy-momentum tensor, describing the density of matter. Where there is no matter, the energy-momentum tensor is zero and the space-time is Ricci-flat; and where matter is present, the Ricci tensor is locally determined by this matter density. The remainder of the curvature, namely that measured by the Weyl conformal curvature tensor $\mathbf{C}$, is governed by differential equations and is thus constrained, in a nonlocal way, by the distribution of matter. In the limit when velocities are small and the fields are weak (in the sense of small gravitational potentials) the theory goes over into the older Newtonian theory. Various small deviations from Newtonian theory have been observationally established, and these are all consistent with the more accurate Einstein theory. In fact, Einstein's theory must now be regarded as an excellently-tested theory of gravity, having no serious rival (at least none for which there is any forseeable prospect of observing a discrepancy with Einstein's theory). The mathematics of Einsteinian manifolds —and including that of the important special case provided by Minkowski geometry—is thus something of especial interest for physics.

3. Twistor geometry and CR-structure. Let us next examine how the geometry of Minkowski space M may be reformulated in a new way (cf. Penrose (1975), (1977); Penrose and Ward (1980); also Wells (1979)). We shall see how this relates M to the theory of boundaries of complex manifolds. An analogous construction for the more general curved Einstein spaces will be considered afterwards.

Choose standard Minkowski coordinates $x = (\tau, \xi, \eta, \zeta) \in \mathbb{R}^4$, for M, where S is given by

$$\begin{aligned} S(x, \hat{x}) &= (\tau - \hat{\tau})^2 - (\xi - \hat{\xi})^2 - (\eta - \hat{\eta})^2 - (\zeta - \hat{\zeta})^2 \\ &= \det\left\{\begin{pmatrix} \tau + \zeta & \xi + i\eta \\ \xi - i\eta & \tau - \zeta \end{pmatrix} - \begin{pmatrix} \hat{\tau} + \hat{\zeta} & \hat{\xi} + i\hat{\eta} \\ \hat{\xi} - i\hat{\eta} & \hat{\tau} - \hat{\zeta} \end{pmatrix}\right\} \end{aligned}$$

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

where throughout this work "i" stands for $\sqrt{-1}$ (and a bar denotes complex conjugation).

Let $(W_0, W_1, W_2, W_3) \in \mathbb{C}^4$ be coordinates for an element W of an associated complex vector space $\mathbb{T}.$ called (dual) twistor space where W and x are said to be *incident* whenever

$$(W_2, W_3) = \frac{1}{i\sqrt{2}}(W_0, W_1)\begin{pmatrix} \tau + \zeta & \xi + i\eta \\ \xi - i\eta & \tau - \zeta \end{pmatrix}. \tag{3.2}$$

This relation can only hold if $\Sigma(W) = 0$, where the $(+,+,-,-)$-signature Hermitian form Σ is defined by

$$\Sigma(W) \equiv W_0\overline{W_2} + W_1\overline{W_3} + W_2\overline{W_0} + W_3\overline{W_1} \tag{3.3}$$

(because the Hermiticity of the (2×2) matrix in (3.2) entails that postmultiplication of (3.2) by the conjugate transpose of (W_0, W_1) yields a purely imaginary result).

The 3-complex-dimensional projective space $\mathbb{P}\mathbb{T}.$, associated with $\mathbb{T}.$, whose points are labelled by the three complex ratios

$$W_0 : W_1 : W_2 : W_3$$

consists of two open complex manifolds $\mathbb{P}\mathbb{T}_+$ and $\mathbb{P}\mathbb{T}_-$, given when $\Sigma(W) > 0$ and $\Sigma(W) < 0$, respectively, together with their common boundary, the 5-real-dimensional manifold $\mathbb{P}\mathbb{T}_0$, given when $\Sigma(W) = 0$. Thus the points of $\mathbb{P}\mathbb{T}.$ for which (3.2) holds for some $x \in M$, all lie on $\mathbb{P}\mathbb{T}_0$. They do not constitute quite the whole of $\mathbb{P}\mathbb{T}_0$, however, since points of the projective line I, given by $W_0 = W_1 = 0$, admit no solution for x in (3.2).

Suppose we choose a fixed

$$\mathbf{W} \in \mathbb{P}\mathbb{T}_0 - I$$

(where from now on I use boldface capital letters such as $\mathbf{W}$ to denote the point of the projective space $\mathbb{P}\mathbb{T}.$, rather than $\mathbb{T}.$ or $\mathbb{C}^4$). Then we can solve (3.2) for the Minkowski point x. The solution is not unique, however, the freedom being given by

$$\begin{pmatrix} \tau + \zeta & \xi + i\eta \\ \xi - i\eta & \tau - \zeta \end{pmatrix} \mapsto \begin{pmatrix} \tau + \zeta & \xi + i\eta \\ \xi - i\eta & \tau - \zeta \end{pmatrix} + k\begin{pmatrix} W_1\overline{W_1} & -W_1\overline{W_0} \\ -W_0\overline{W_1} & W_0\overline{W_0} \end{pmatrix}$$

for arbitrary real k. Note that the final matrix on the right has rank unity and so represents, by (3.1), a *null* (i.e. zero Minkowski length) vector in M. Thus, the points x incident with the given W constitute a *null geodesic* (straight line) in M which, without confusion, we can also label by $\mathbf{W}$. Indeed, we may think of $\mathbb{P}\mathbb{T}_0 - I$ *as* the space of null geodesics $\mathbf{W}$ in M. Our construction has shown how this space may be imbedded as a 5-dimensional real submanifold of a complex projective 3-space.

To interpret a point $x \in M$, conversely in terms of $\mathbb{P}\mathbb{T}.$, we fix x in (3.2) and allow $\mathbf{W}$ to vary. This gives us a 2-complex-dimensional linear subspace in $\mathbb{T}.$, i.e.

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

a complex projective line in $\mathbb{PT}$.. Clearly this line lies in $\mathbb{PT}_0 - I$ and, moreover, all lines lying entirely in $\mathbb{PT}_0 - I$ arise in this way from points in M. Thus

the projective lines lying entirely in $\mathbb{PT}_0 - I$
represent the points of Minkowski space M.

Note that if we allow the coordinates τ, ξ, η, ζ to become complex i.e. x to become a point of the *complexification* $\mathbb{C}M$ of M, then the above construction yields a projective line lying entirely in $\mathbb{PT}. - I$. We may also consider projective lines in $\mathbb{PT}_0$ [resp. $\mathbb{PT}.$] which meet I. These provide "points at infinity" for M [resp. $\mathbb{C}M$], the totality of all lines in $\mathbb{PT}_0$ [resp. $\mathbb{PT}.$] describing the standard conformal compactification $M^\#$ [resp. $\mathbb{C}M^\#$] of M [resp. $\mathbb{C}M$].

The intrinsic structure of projective lines in $\mathbb{PT}_0 - I$ can be illustrated in a very graphic "physical" way. The family of light rays **W** through (i.e. incident with) a fixed point $x \in M$ represents the family of points on the corresponding complex projective line in $\mathbb{PT}_0 - I$. A complex projective line has the topology S^2 and, moreover, the holomorphic structure of a Riemann sphere. Now imagine an observer situated at x. His field of vision will be represented by the light rays through x, i.e. by the family of **W**'s under consideration. Clearly the topology of the observer's entire field of vision is indeed S^2. But, more subtly, the *holomorphic* structure is also relevant. This shows up in the allowed transformations between various observers with different velocities, all of whom pass through the same space-time event x. Their fields of vision are related to one another, at x, by transformations which preserve the holomorphic structure of the Riemann sphere, i.e. their fields of vision are *conformally* related to one another. This is a well-known property of Lorentz transformations (Penrose (1959), Terrell (1959)). Indeed, the connected component of the *Lorentz group* may be regarded as the group of all *holomorphic self-transformations* of this Riemann sphere.

By virtue of its imbedding in the complex manifold $\mathbb{PT}$., the entire (real) hypersurface $\mathbb{PT}_0$ inherits further holomorphic structure. This structure, referred to as a (maximal[1]) CR-structure falls short by just one (real) dimension of defining $\mathbb{PT}_0$ as a complex manifold. We shall see the physical interpretation of this shortly.

The structure of a (maximal) CR-manifold $\mathcal{V}$ of, say, dimension $2n + 1$ provides that in the tangent space T_p, at each point $p \in \mathcal{V}$, a $2n$-real-dimensional subspace $H_p \subset T_p$ is singled out, referred to as the *holomorphic tangent space* at p. The space H_p is to be regarded as a *complex* vector space of n dimensions, spanned by complex vectors

$$\mathbf{z}_1 = \mathbf{x}_1 + i\mathbf{y}_1, \ldots, \mathbf{z}_n = \mathbf{x}_n + i\mathbf{y}_n.$$

[1]Some authors use the term "CR-structure" also in the extended sense of the structure induced on any real submanifold of a complex manifold, not necessarily a hypersurface. The word "maximal" is to emphasize that I am concerned only with the hypersurface dimensionality here.

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

The real-linear operator J, satisfying $J^2 = -1$, is determined by its action on the related real basis $\mathbf{x}_1,\ldots,\mathbf{x}_n, \mathbf{y}_1,\ldots,\mathbf{y}_n$ for H_p by

$$J\mathbf{x}_r = -\mathbf{y}_r, \qquad J\mathbf{y}_r = \mathbf{x}_r \qquad (r = 1,\ldots,n) \tag{3.4}$$

so that

$$J\mathbf{z}_r = i\mathbf{z}_r \qquad (r = 1,\ldots,n). \tag{3.5}$$

It is the operator J that defines the complex structure of H_p when we regard H_p as a real vector space. One further tangent vector $\mathbf{u} \in T_p$, independent of $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $\mathbf{y}_1,\ldots,\mathbf{y}_n$, is needed to complete a basis for the whole of T_p. We assume that in some neighbourhood $\mathcal{U}_p$ of p in $\mathcal{V}$ a smooth choice of such basis vectors is made and we define the *integrability relations* for the CR-structure of $\mathcal{V}$, defined by J, to be the Lie bracket relation

$$[\mathbf{z}_r, \mathbf{z}_{r'}] = \text{complex linear combinations of } \mathbf{z}\text{'s} \tag{3.6}$$

$(r, r' = 1,\ldots,n)$. It is clear that this integrability relation is independent of the particular choice of basis satisfying (3.5), and depends only on the choice of J and the holomorphic subspaces.

Note that what distinguishes the structure of $\mathcal{V}$ from that of a complex manifold (of dimension n) is merely the presence of the extra real dimension defined by the extra real basis vector $\mathbf{u}$.

In the case of a *complex manifold* we know from the Newlander-Nirenberg (1957) theorem that integrability relations like (3.6) are all that are required. No analyticity assumptions are needed to ensure that holomorphic coordinates can be locally introduced, the derivatives with respect to which providing complex tangent vectors compatible with the given complex structure J in the sense of (3.5). The analogous property for a (maximal) CR-manifold $\mathcal{V}$ of real dimension $2n + 1$ would be the existence of a complex $(n + 1)$-manifold $\mathcal{C}$, with complex structure J, in which $\mathcal{V}$ can be realized as a real hypersurface, the CR-structure on $\mathcal{V}$ being that induced by J. However, counterexamples given by Nirenberg [(1973), (1974)] show that the integrability relations (3.6) are insufficient to ensure that the CR-structure of $\mathcal{V}$ can be so realized in all cases.

It should be made clear that the difficulty does not lie in a possible omission of necessary additional integrability relations of the usual (differential) kind. Indeed, if the CR-structure is taken to be real-analytic, then the integrability relations (3.6) are actually sufficient to ensure that $\mathcal{V}$ can be realized in the above way. But for a C^∞ CR-structure this need not be so. In this respect, the problem of realizing CR-structures is analogous to that of the Lewy insoluble differential equation (1957). We shall see later how this can be made more explicit.

Let us return to $\mathbb{PT}_0$, which has a CR-structure explicitly realized by its imbedding in the complex manifold $\mathbb{PT}$.. Here $n = 2$. Thus, at any point $\mathbf{W}$ of $\mathbb{PT}_0$ the 5-real-dimensional tangent space $T_{\mathbf{W}}$ has a 2-complex-dimensional holomorphic subspace $H_{\mathbf{W}}$ whose complete structure is defined by a certain operator J. Let us first identify the location of $H_{\mathbf{W}}$ in physical terms. The point $\mathbf{W} \in \mathbb{PT}_0$ is interpreted as a light ray in M. We may picture this by fixing the "time" τ to

take a specific value, say $\tau = 0$, so our picture becomes that of ordinary Euclidean 3-space U. Then the "photon" described by $\mathbf{W}$ is pictured as a *point* in U (its location at $\tau = 0$) together with a *direction* at that point defining its velocity (the magnitude of the velocity being unity: the "speed of light"), i.e. we represent $\mathbf{W}$ by a unit vector $\mathbf{w}$ at some point q of U. To picture $T_{\mathbf{W}}$ we envisage making a small displacement of q and $\mathbf{w}$. If this displacement is such that q is moved in a direction orthogonal to the direction of $\mathbf{w}$, then we get something in the subspace $H_{\mathbf{W}}$. We refer to these slightly displaced light rays as being *abreast* with one another. The subspace $H_{\mathbf{W}}$ divides $T_{\mathbf{W}}$ into two remaining pieces, representing light rays that are slightly ahead or lag slightly behind the original ray $\mathbf{W}$. These properties are easily seen to be independent of the "time" τ, i.e. independent of the particular Euclidean (spacelike) hyperplane U that is chosen for our description.

To picture the effect of J, imagine a 2-plane element π, at q, which is orthogonal to the direction $\mathbf{w}$. We are concerned only with displacements of q and $\mathbf{w}$ for which q is moved within π, since these correspond to real vectors in $H_{\mathbf{W}}$ (i.e. "abreast" displacements of $\mathbf{W}$). Such displacements are represented by pairs of vectors $(\mathbf{r}, \mathbf{v})$ in π, where $\mathbf{r}$ gives the displacement of q and $\mathbf{v}$ measures the change in $\mathbf{w}$ (clearly also orthogonal to $\mathbf{w}$, since $\mathbf{w}$ is a unit vector). If we keep the light ray $\mathbf{W}$ and the "neighbouring" ray to which it is displaced both fixed, but vary the time τ, we easily find that the functions $\mathbf{r}(\tau)$, $\mathbf{v}(\tau)$ are related by

$$d\mathbf{r}/d\tau = \mathbf{v}, \qquad d\mathbf{v}/d\tau = 0. \tag{3.7}$$

Now the effect of J turns out to be simply to *rotate both* $\mathbf{r}$ *and* $\mathbf{v}$ *through a right angle in the plane* π, in a left-handed sense about the direction of $\mathbf{w}$. Clearly (3.7) is invariant under this operation, so J applies to the displacements of light rays in their entirety and not simply to light rays relative to particular points on them. (Moreover this operation is easily seen to be independent of the slope of the hyperplane U.) Thus, the CR-structure of $\mathbb{P}\mathcal{T}_0$ is seen to have direct interpretation in terms of physical space-time geometry.

Let us next consider what happens when Minkowski space M is replaced by a more general Einsteinian space-time manifold $\mathcal{M}$. We shall suppose, for simplicity that $\mathcal{M}$, with its pseudometric $\mathbf{g}$, is *globally hyperbolic* (Leray (1952)). According to a result of Geroch (1970) this property can be stated as the existence of a spacelike hypersurface $\mathcal{U}$ in $\mathcal{M}$ which intersects every null geodesic (light ray) in $\mathcal{M}$ in precisely one point. ("Spacelike", in this context, means that its normal vectors are everywhere timelike, i.e. that its induced metric is everywhere negative definite.)

It turns out that the space $\mathbb{P}\mathcal{T}_0$ of null geodesics in $\mathcal{M}$ acquires a CR-structure relative to the hypersurface $\mathcal{U}$, where the holomorphic subspaces to the tangent spaces in $\mathbb{P}\mathcal{T}_0$ and the operator J, now denoted $J_{\mathcal{U}}$, are defined essentially as before. Thus we may represent any $\mathbf{W} \in \mathbb{P}\mathcal{T}_0$ by the point $q \in \mathcal{U}$ at which $\mathcal{U}$ intersects the null geodesic $\mathbf{W}$, together with the unit vector $\mathbf{w}$ in $\mathcal{U}$ at q in the direction of the orthogonal projection into $\mathcal{U}$ of the future-pointing null direction

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society,* Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

of **W**. The holomorphic tangent space to $\mathbb{P}\mathcal{T}_0$ at **W** is provided by those displacements of **W** for which q moves in a direction orthogonal to **w**, and the (real) vectors in this space therefore correspond to pairs $(\mathbf{r}, \mathbf{v})$, as before, where **r** and **v** are (real) tangent vectors to $\mathcal{U}$ at q lying in the 2-plane element π orthogonal to **w**. The effect of $J_{\mathcal{U}}$ is, as before, to rotate **r** and **v** through a right angle, in a left-handed sense about **w**. It then turns out that the integrability relations (3.6) for $J_{\mathcal{U}}$ are automatically satisfied (Penrose (1975), LeBrun (1980)) and the required CR-structure, denoted $\mathrm{CR}_{\mathcal{U}}$, is thereby obtained.

However, unlike the CR-structure for $\mathbb{P}\mathbb{T}_0$, $\mathrm{CR}_{\mathcal{U}}$ will in general depend upon the location of $\mathcal{U}$ in $\mathcal{M}$. The holomorphic tangent spaces, regarded as real vector subspaces of the tangent spaces to $\mathbb{P}\mathcal{T}_0$, are in fact independent of $\mathcal{U}$, but their complex structures, as defined by $J_{\mathcal{U}}$ will generally vary. These properties follow from the Jacobi equation, which suitably replaces the second of equations (3.7). Only when the metric **g** of $\mathcal{M}$ is conformally flat will $\mathrm{CR}_{\mathcal{U}}$ be completely independent of $\mathcal{U}$.

By way of clarification of this point it may be remarked that even if $\mathcal{M}$ is *stationary* (i.e. time-independent, in the sense of possessing a timelike killing vector), so that isometries of $\mathcal{M}$ exist carrying $\mathcal{U}$ into a succession of geometrically equivalent spacelike hypersurfaces $\mathcal{U}', \mathcal{U}'', \ldots, \subset \mathcal{M}$, the structures $\mathrm{CR}_{\mathcal{U}}$, $\mathrm{CR}_{\mathcal{U}'}$, $\mathrm{CR}_{\mathcal{U}''}, \ldots$ on $\mathbb{P}\mathcal{T}_0$ will generally be distinct. This is because a particular point **W** of $\mathbb{P}\mathcal{T}_0$ represents a nonstationary object (namely a null geodesic) in $\mathcal{M}$, so it is related differently to each of $\mathcal{U}$, $\mathcal{U}'$, $\mathcal{U}'', \ldots$. However, in this case the isometries of $\mathcal{M}$ will carry **W** into null geodesics $\mathbf{W}', \mathbf{W}'', \ldots$, so $\mathrm{CR}_{\mathcal{U}}$ at **W** will agree with $\mathrm{CR}_{\mathcal{U}'}$ at $\mathbf{W}'$ and with $\mathrm{CR}_{\mathcal{U}''}$ at $\mathbf{W}''$, etc. Thus the structures $\mathrm{CR}_{\mathcal{U}}, \mathrm{CR}_{\mathcal{U}'}, \ldots$ will in this case be *intrinsically* identical even though they represent generally distinct CR-structures on the given space $\mathbb{P}\mathcal{T}_0$.

The fact that CR-structures (maximal, and of given dimension $2n + 1 > 0$) *can* be locally *distinct* from one another is, in effect, an observation that dates back to an important paper of Poincaré (1907). What Poincaré actually showed was that the Riemann mapping theorem, which states that any smoothly bounded simply-connected region $\mathcal{D}$ in the Argand plane $\mathbb{C}^1$ is holomorphically identical with the unit disc, has no direct analogue in higher complex dimension. Roughly, the argument is to show that the (smooth) boundary of a region in, say, $\mathbb{C}^2$ contains intrinsic holomorphically invariant information about its "shape". (This does not occur for a smooth curve in $\mathbb{C}^1$, as the Riemann mapping theorem shows.) The gist of Poincaré's argument can be simplified to the following "physicists proof". Consider the freedom involved in specifying the (smooth) real hypersurface boundary of a region in $\mathbb{C}^2$. This is provided by one real function of three real variables (e.g. $\mathrm{Im}\,\zeta_1$ in terms of $\mathrm{Re}\,\zeta_1$, $\mathrm{Re}\,\zeta_2$, $\mathrm{Im}\,\zeta_2$). For the *intrinsic* structure of this boundary, we must factor out by the freedom provided by the allowed local holomorphic maps of $\mathbb{C}^2$ to itself. This is provided (locally) by two holomorphic functions of two complex variables. But a holomorphic function is determined by its (analytic) values on any real environment, e.g. on its values where ζ_1 and ζ_2 are

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society,* Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

real. The real and imaginary parts there are each, in effect, independent real analytic functions, so the freedom to be factored out by is that of four real functions of two real variables. The amount of intrinsic freedom in the structure of the boundary is therefore

$$\frac{\text{1 real function of 3 real variables}}{\text{4 real functions of 2 real variables}}.$$

However, any finite number of functions of two variables must be regarded as "peanuts" in the context of free functions of three variables; i.e. it is completely swamped by the three-variables' worth of freedom and makes no contribution net count. Thus there must be "intrinsic invariants" of the boundary shape whose functional freedom is such as to be dependent upon three real variables. More generally, for a boundary in $\mathbb{C}^{n+1}$ (and so for a $(2n+1)$-dimensional CR-manifold) the freedom for the invariants is functions of $2n+1$ variables ($n > 0$).

The detailed form of these invariants was investigated thoroughly by É. Cartan (1932), Tanaka (1962) and Chern & Moser (1974). This study may be regarded as the analogue, for CR-manifolds, of the study of invariants (and covariant tensorial objects) for ordinary Riemannian geometry.

It turns out that the structures $\mathrm{CR}_{\mathfrak{N}}$ that we have just been considering do in fact differ, in general, from the original CR-structure of $\mathbb{PT}_0$ that arose in relation to Minkowski space. There is the further point that by choosing $\mathfrak{M}$ and $\mathfrak{N}$ to be suitably nonanalytic, the structure $\mathrm{CR}_{\mathfrak{N}}$ can be made to be *nonrealizable* as a real hypersurface in a complex 3-manifold, analogously to the Nirenberg counterexamples referred to earlier. This was first demonstrated by C. R. LeBrun (1980), (1982). Some of the relevant ideas and a slightly different, but related, class of counterexamples had been put forward earlier by G. A. J. Sparling (and involved a suggestion by the present author). A development of Sparling's original line of thinking will be presented in the next section.

4. Boundaries of cohomology classes and complex manifolds. The first and most important invariant of CR-structure is the signature of the Levi form. One way of defining the Levi form for a $(2n+1)$-dimensional CR-manifold $\mathcal{V}$ is as follows. Suppose $\mathcal{V}$ is given, locally, as a real hypersurface

$$\Sigma(\zeta_0,\ldots,\zeta_n) = 0$$

in a complex n-manifold $\mathcal{C}$, where Σ is a real (at least C^2) function of local holomorphic coordinates $\zeta_0,\ldots,\zeta_n$ for $\mathcal{C}$. Then the Hermitian form whose matrix is

$$\frac{\partial\Sigma}{\partial\zeta_j\partial\bar{\zeta}_k},$$

but restricted to the holomorphic tangent space H_p, defines the Levi form at p. Alternatively an entirely intrinsic definition can be given which does not depend upon an imbedding in a complex $(n+1)$-manifold and so applies to nonrealizable CR-structures. Let the complex basis vectors $\mathbf{z}_1,\ldots,\mathbf{z}_n$ and real basis vector $\mathbf{u}$

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

be as in §3 (cf. (3.5), (3.6)). Then the Lie brackets

$$[\mathbf{z}_j, \bar{\mathbf{z}}_k] = iL_{jk}\mathbf{u} + \text{terms in } \mathbf{z}\text{'s and } \bar{\mathbf{z}}\text{'s} \tag{4.1}$$

serve to define the matrix L_{jk} of the Levi form intrinsically.

Though the matrix of the Levi form is clearly not an invariant, its signature—by which I mean the number of plus signs, minus signs and zeros that it acquires when unitarily reduced to diagonal form—is invariant. A remark about the *sign* of the Levi form should also be made here. Suppose we are concerned with the (smooth) boundary $\mathcal{V}$ of a complex manifold; then we should take that complex manifold on the side on which Σ is negative. Thus, for example, for the unit ball in $\mathbb{C}^{n+1}$, we can take

$$\Sigma = \zeta_0\bar{\zeta}_0 + \cdots + \zeta_n\bar{\zeta}_n - 1$$

and the Levi form is positive definite $(+ + \cdots +)$, this being the situation in which the manifold is referred to as (holomorphically) strictly *pseudoconvex*. Correspondingly, if we use the intrinsic method (4.1), we choose the vector $\mathbf{u}$ so that $J\mathbf{u}$ points *outwards* away from the manifold whose boundary we are considering. Generally, if the Levi form signature has j plus signs and k minus signs at a point of $\mathcal{V}$, I shall say that this manifold has j degrees of pseudoconvexity and k degrees of pseudoconcavity at p.

An important result, effectively due to Hans Lewy (1956) (and Bochner (1943); cf. also Hörmander (1966)) states if f is a CR-function on $\mathcal{V}$—which means that

$$\bar{z}_1(f) = 0, \ldots, \bar{z}_n(f) = 0$$

—where we assume that $\mathcal{V}$ is realizable as a real hypersurface in a complex manifold $\mathcal{C}$, then f extends (locally) as a holomorphic function on any side of $\mathcal{V}$ having at least one degree of pseudoconvexity.

It should be remarked that any holomorphic function in $\mathcal{C}$ clearly restricts to a CR-function on $\mathcal{V}$. Moreover, any holomorphic function on one side of $\mathcal{V}$ which attains smooth values at $\mathcal{V}$ itself will also restrict to a CR-function on $\mathcal{V}$. The (Lewy) result just stated shows that a holomorphic function defined on a smoothly bounded region $\mathcal{D}$ of a complex manifold is always locally extendible beyond that boundary at points where $\mathcal{D}$ has at least one degree of pseudoconcavity. Thus smoothly bounded domains of holomorphy, on which locally inextendible holomorphic functions exist, have no points with any degree of pseudoconcavity—i.e. the Levi signature is always of the form $(+ \cdots + 0 \cdots 0)$.

Let us now consider the 5-dimensional real hypersurface $\mathbb{PT}_0$ as a common boundary between $\mathbb{PT}_-$ and $\mathbb{PT}_+$. By simple direct calculation from (3.3), we find that the Levi form signature is $(+,-)$, while for the 7-dimensional real hypersurface $\mathbb{T}_0$ as a boundary between $\mathbb{T}_-$ and $\mathbb{T}_+$ the signature is $(+,-,0)$. (The zero arises from the holomorphic direction corresponding to complex rescalings $W \mapsto \lambda W$, $\lambda \in \mathbb{C} - \{0\}$.) From the results just stated we see that any CR-function defined in some open neighbourhood of a point in $\mathbb{PT}_0$ [resp. $\mathbb{T}_0$] will extend to a holomorphic function in a neighbourhood in $\mathbb{PT}_.$ [resp. $\mathbb{T}_.$].

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

It turns out, however, that the most directly physical objects on these spaces are not holomorphic or CR-functions but objects of the next degree of abstraction, namely *first sheaf cohomology classes*. The simplest example of the physical interpretation of such a cohomology class can be illustrated as follows. Consider some suitable open neighbourhood $\mathcal{Q}$, in $\mathbb{C}M$, of a point $x \in \mathbb{C}M$. According to the discussion of §3, each point of $\mathcal{Q}$ will be represented by a projective line in $\mathbb{P}\mathbb{T}$., so as the point varies throughout $\mathcal{Q}$ the corresponding line sweeps out an open region $\mathcal{L} \subset \mathbb{P}\mathbb{T}$.. We shall be concerned with an element

$$\Phi \in H^1(\mathcal{L}, \mathcal{O}(-2))$$

where, generally, $\mathcal{O}(r)$ denotes the sheaf of germs of holomorphic functions on $\mathbb{P}\mathbb{T}$. "twisted by r"—i.e. represented in terms of holomorphic functions on $\mathbb{T}$. which are *homogeneous of degree r*. Under fairly liberal restrictions on the nature of $\mathcal{Q}$ (cf. Eastwood, Penrose and Wells (1981) for details) we find that Φ represents a solution ϕ of the wave equation

$$\Box\phi \equiv \left(\frac{\partial^2}{\partial\tau^2} - \frac{\partial^2}{\partial\xi^2} - \frac{\partial^2}{\partial\eta^2} - \frac{\partial^2}{\partial\zeta^2}\right)\phi = 0 \tag{4.2}$$

in $\mathcal{Q}$.

Suppose we have the simple situation in which Φ can be described by a representative Čech cocycle Φ_{12} $(= -\Phi_{21})$ for the 2-set cover $\mathcal{U}_1$, $\mathcal{U}_2$ of $\mathcal{L}$, so Φ_{12} is simply a holomorphic function (taken as homogeneous of degree -2 in $W_0, \ldots, W_3$) on $\mathcal{U}_{12} = \mathcal{U}_1 \cap \mathcal{U}_2$, where $\mathcal{L} \subset \mathcal{U}_1 \cup \mathcal{U}_2$. To obtain ϕ, we then perform the contour integral

$$\phi(x) = \frac{1}{2\pi i}\oint_{\Xi} \Phi_{12}(W)\Delta W \tag{4.3}$$

with

$$\Delta W = W_0 dW_1 - W_1 dW_0 \tag{4.4}$$

where the 1-dimensional closed contour Ξ (in $\mathbb{P}\mathbb{T}$.) lies in the intersection of $\mathcal{U}_{12}$ with the projective line $L_x \subset \mathbb{P}\mathbb{T}$. representing $x = (\tau, \xi, \eta, \zeta)$. Note that since we are integrating over points of $\mathbb{P}\mathbb{T}$. *incident* with x, we can use the explicit expression (3.2) to represent W_2 and W_3 in terms of W_0 and W_1. Moreover, because of the -2 homogeneity of Φ_{12} and the form of the expression (4.4), the exterior derivative of the integrand in (4.3) (for W incident with x) vanishes. Thus (4.3) is a genuine contour integral for fixed x, depending only on the homology class of the contour within $L_x \cap \mathcal{U}_{12}$. Simple direct calculation (using (3.2)) shows that (4.2) is indeed satisfied as required.

We could also envisage a more complicated cover $\mathcal{U}_1, \ldots, \mathcal{U}_r$ of $\mathcal{L}$, with Φ given a Čech description with respect to it by the family of holomorphic (homogeneous of degree -2) functions

$$\Phi_{jk} = -\Phi_{kj} \quad \text{on } \mathcal{U}_{jk} = \mathcal{U}_j \cap \mathcal{U}_k \tag{4.5}$$

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

subject to

$$\Phi_{jk} + \Phi_{kl} + \Phi_{lj} = 0 \quad \text{on } \mathcal{U}_{jkl} = \mathcal{U}_j \cap \mathcal{U}_k \cap \mathcal{U}_l. \tag{4.6}$$

The expression for $\phi(x)$ is given, essentially as before, by (4.3), but now the contour Ξ is a branched one, consisting of various segments $\Xi_{jk} \subset \mathcal{U}_{jk}$, over which Φ_{jk} is to be integrated, with endpoints $\mathbf{W}_{jkl} \in \mathcal{U}_{jkl}$, and we sum over the contributions from these various Φ_{jk}. In fact (for appropriate $\mathcal{Q}$) the general solution of (4.2) can be obtained in this way (cf. Eastwood, Penrose and Wells (1981), Penrose and Ward (1980)).

Similarly Maxwell's free-space equations can be solved in terms of elements

$$(\Theta, \Psi) \in H^1(\mathbb{L}, \mathcal{O}(-4)) \oplus H^1(\mathbb{L}, \mathcal{O}(0)). \tag{4.7}$$

The components of the Maxwell field tensor, in the (τ, ξ, η, ζ)-coordinate system, are given by

$$\begin{pmatrix} 0 & \theta_0 - \theta_2 + \psi_0 - \psi_2 & i(\theta_0 + \theta_2 - \psi_0 - \psi_2) & -2(\theta_1 + \psi_1) \\ \theta_2 - \theta_0 + \psi_2 - \psi_0 & 0 & 2i(\theta_1 - \psi_1) & -\theta_0 - \theta_2 - \psi_0 - \psi_2 \\ i(-\theta_0 - \theta_2 + \psi_0 + \psi_2) & 2i(-\theta_1 + \psi_1) & 0 & i(\theta_2 - \theta_0 - \psi_2 + \psi_0) \\ 2(\theta_1 + \psi_1) & \theta_0 + \theta_2 + \psi_0 + \psi_2 & i(\theta_0 - \theta_2 - \psi_0 + \psi_2) & 0 \end{pmatrix}$$

where

$$\begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix} = \frac{1}{2\pi i} \oint_\Xi \begin{pmatrix} W_0^2 \\ W_0 W_1 \\ W_1^2 \end{pmatrix} \Theta(W) \Delta W, \qquad \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \end{pmatrix} = \frac{1}{2\pi i} \oint_\Xi \begin{pmatrix} \partial^2/\partial W_2^2 \\ \partial^2/\partial W_2 \partial W_3 \\ \partial^2/\partial W_3^2 \end{pmatrix} \Psi(W) \Delta W$$

(cf. Penrose (1969), (1977)), the notation $\Theta(W)$, $\Psi(W)$ indicating the appropriate $\Theta_{jk}(W)$, $\Psi_{jk}(W)$ to be integrated over Ξ and summed as before. The quantities $\theta_0, \theta_1, \theta_2$ define the anti-self-dual part of the Maxwell field and ψ_0, ψ_1, ψ_2 define the self-dual part. For a *real* Maxwell field we require

$$\theta_0 = \bar{\psi}_0, \qquad \theta_1 = \bar{\psi}_1, \qquad \theta_2 = \bar{\psi}_2,$$

at real points x (and, at complex points, $\theta_0(x) = \bar{\psi}_0(\bar{x})$, etc.) so we can use *either* $H^1(\mathbb{L}, \mathcal{O}(-4))$ *or* $H^1(\mathbb{L}, \mathcal{O}(0))$ to specify such a field, employing complex conjugation in place of using the other.

Recall that the Levi form signature for a smoothly bounded domain of holomorphy—i.e. natural (smooth) boundary of the domain of some holomorphic *function*—is always of the form $(+ \cdots + 0 \cdots 0)$. But a function is a *zeroth* sheaf cohomology element (an H^0), whereas we are now more concerned with *first* cohomology elements (H^1's). There is an analogous concept to a domain of holomorphy for such objects, namely a smoothly bounded region $\mathcal{D}$ on which is defined some holomorphic H^1 (i.e. an H^1 for a coherent analytic sheaf) which does not locally extend beyond the boundary $\partial\mathcal{D}$. It follows from a theorem of Andreotti and Hill (1972) that, at each point of $\partial\mathcal{D}$, $\mathcal{D}$ either has precisely *one* degree of pseudoconcavity (i.e. Levi form signature of the form $(+ \cdots + 0 \cdots 0 -)$) or else is pseudoconvex with at least one zero in the Levi signature (i.e.

$(+\cdots+0\cdots00)$). Referring to the number of zeros in the Levi signature as the number, d, of *degrees of holomorphic flatness* we find, generally, that a smoothly bounded region on which is defined a locally inextendible holomorphic H^r has a number, q, of degrees of pseudoconcavity satisfying, at each point of its boundary,

$$r \geqslant q \geqslant r - d. \tag{4.8}$$

We have seen that for ordinary functions (i.e. H^0's), we could refer to the concept of CR-function on a CR-manifold $\mathcal{V}$, realized as a hypersurface in a complex manifold $\mathcal{C}$, and ask whether holomorphic extendibility on one side or the other is locally possible. The appropriate concept for H^r's can be stated in terms of the $\bar{\partial}_b$-cohomology of Andreotti & Hill (1972) and Hill & McKichan (1977); cf. also Polking & Wells (1976), Hill (1973). We find that an appropriate concept of a "CR-H^r" is *not* taken as the restriction to $\mathcal{V}$ of a holomorphic H^r in $\mathcal{C}$ but may be defined, rather, in terms of the $\hat{H}^r$ provided by an exact sequence

$$\cdots \to H^r(\mathcal{C}) \to \bigoplus_{H^r(\mathcal{C}^-)}^{H^r(\mathcal{C}^+)} \to \hat{H}^r(\mathcal{V}) \to H^{r+1}(\mathcal{C}) \to \bigoplus_{H^{r+1}(\mathcal{C}^-)}^{H^{r+1}(\mathcal{C}^+)} \to \cdots \tag{4.9}$$

where $\mathcal{V}$ separates $\mathcal{C}$ into the two pieces $\mathcal{C}^-$, $\mathcal{C}^+$ (so $\mathcal{C} = \mathcal{C}^- \cup \mathcal{C}^+ \cup \mathcal{V}$). Here $\hat{H}^r(\mathcal{V})$ denotes the required space of rth *hyperfunctional* $\bar{\partial}_b$-cohomology. (See Sato (1959–60), Kashiwara (1979).) The sheaf for each H is taken to be that of local cross-sections of some given holomorphic vector bundle over $\mathcal{C}$, and that for $\hat{H}$ to correspond in an appropriate sense to the restriction of this bundle to $\mathcal{V}$.

If we can arrange $H^{r+1}(\mathcal{C}) = 0$, which will be the case, for example, when $\mathcal{C}$ is *Stein* (i.e. a domain of holomorphy), then we obtain from (4.9)

$$\hat{H}^r(\mathcal{V}) = \frac{H^r(\mathcal{C}^-) \oplus H^r(\mathcal{C}^+)}{H^r(\mathcal{C})} \tag{4.10}$$

—and if $\mathcal{C}$ is Stein we can ignore the $H^r(\mathcal{C})$ $(= 0)$ when $r > 0$. We can loosely interpret the elements of $\hat{H}^r(\mathcal{V})$, in this description, as the space of possible "jumps" in H^r as we cross from one side of $\mathcal{V}$ to the other. In the particular case $r = n = 0$ (where $\mathcal{C}$ has complex dimension $n + 1$), with $\mathcal{V}$ being (a portion of) the real axis in $\mathbb{C}$, (4.10) provides the standard definition of 1-dimensional hyperfunctions on $\mathbb{R}$ (Sato (1959–60), Kashiwara (1979)). Hyperfunctions are generalizations of distributions and so they include, as special cases, ordinary C^∞ functions. The C^∞ CR-functions on $\mathcal{V}$ can thus be taken to be *particular* elements of $\hat{H}^0(\mathcal{V})$.

The concept of local (Lewy) extendibility of C^∞ CR-functions on $\mathcal{V}$, into regions of $\mathcal{C}$ locally bounded by $\mathcal{V}$ wherever there is at least one degree of pseudoconcavity, applies also to CR-hyperfunctions (cf. Hill and MacKichan (1977)). Moreover, there is a corresponding result for elements of $\hat{H}^r(\mathcal{V})$ generally: local holomorphic extendibility into the complex manifold on one side of $\mathcal{V}$ can always be achieved, in the case of a nondegenerate Levi form, whenever the number, q, of degrees of pseudoconcavity is not exactly r. More generally, if there are d degrees of holomorphic flatness, then local extendibility can always be

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

achieved wherever

$$q > r \quad \text{or} \quad q + d < r, \tag{4.11}$$

which is the negation of (4.8).

Note that if both $\mathcal{C}^+$ and $\mathcal{C}^-$ have some pseudoconcavity everywhere along $\mathcal{V}$, as is the case when $\mathcal{V} = \mathbb{P}\mathbb{T}_0$ or $\mathbb{T}_0$, then every CR-function (whether C^∞ or hyperfunctional) extends to both sides of $\mathcal{V}$ and is therefore always analytic. This shows that we cannot simply use CR-functions in a straightforward Čech approach to the CR-cohomology of such a $\mathcal{V}$, since nonanalytic elements of H^r will always occur for some $r > 0$. Indeed, when $\mathcal{V} = \mathbb{P}\mathbb{T}_0$ or $\mathbb{T}_0$ we find that $\hat{H}^1(\mathcal{V})$ contains nonanalytic elements. In fact, solutions of the wave equation (4.2) or of Maxwell's equations on (regions in) *real* Minkowski space M can be represented by elements of an $\hat{H}^1(\mathcal{V})$ (respectively $\hat{H}^1(\mathbb{P}\mathbb{T}_0, \mathcal{O}(-2))$ and $\hat{H}^1(\mathbb{P}\mathbb{T}_0, \mathcal{O}(-4) \oplus \mathcal{O}(0))$) and these fields certainly need not be analytic (cf. Wells (1981), Bailey, Ehrenpreis and Wells (1982)). Thus the direct Čech approach, with CR-functions on $\mathbb{P}\mathbb{T}_0$, does not apply and a method such as using the sequence (4.9) or an explicit $\bar{\partial}_b$ approach is called for.[2]

This contrast between a direct Čech approach and such as the $\bar{\partial}_b$ method has its analogue in a related contrast between CR-structures which are realizable in terms of imbeddability in complex manifolds and those which are not. Let us consider two possible approaches to defining a complex manifold. In the first (method $\check{C}$), we consider the manifold to be built up from open patches on $\mathbb{C}^{n+1}$, a family of holomorphic transition functions F_{jk} being provided to "glue" these patches together, where the F_{jk} are subject to certain nonlinear relations closely analogous to (4.5) and (4.6). Likewise, the equivalence between two such complex manifolds so constructed is closely analogous to the condition for two Čech 1-cocycles Φ and Ψ to differ by a coboundary

$$\Phi_{jk} - \Psi_{jk} = \Lambda_k - \Lambda_j \quad \text{on } \mathcal{U}_{jk}$$

where Λ_j is holomorphic on $\mathcal{U}_j$. Here, the Λ_j are analogous to coordinate transformations on the individual patches. Finally, the concepts of taking refinements and direct limits are common both to Čech cohomology and to this method of complex-manifold building. Thus, method $\check{C}$ presents a complex manifold as a kind of nonlinear holomorphic H^1, presented according to a Čech prescription.

In the second method (method $\bar{\partial}$) the manifold is given first as a real manifold of dimension $2(n + 1)$. An operator J (with $J^2 = -1$) is defined to act in the tangent spaces and to satisfy an integrability condition like (3.6) with (3.5). When phrased suitably, method $\bar{\partial}$ resembles a nonlinear version of the Dolbeault ($\bar{\partial}$) approach to the definition of a holomorphic H^1.

---

[2] There are also other (essentially equivalent) methods of defining the concept of a (say C^∞) CR $- H^R$ on $\mathcal{V}$, which can provide the desired nonanalytic cohomology. One of these is to define the cohomology in terms of extensions of CR-vector-bundles over $\mathcal{V}$.

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society,* Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

Because of the Newlander-Nirenberg theorem, these two methods of building a complex manifold are equivalent—and, likewise, these two approaches to constructing a holomorphic H^1 are equivalent. However, we have just seen that there is an inequivalence between the methods of defining the H^1 cohomology at the *boundaries* of complex manifolds. Such considerations led C. D. Hill to suggest, a number of years ago, that there ought to be a *difference between the Čech-method and $\bar{\partial}$-method of defining a complex manifold-with-boundary*, and that this difference should manifest itself when the manifold has one degree of pseudoconcavity at its boundary (or perhaps none, in cases where the Levi form is degenerate). The Čech-method would be as before, except that the patches would now have specified *boundary* portions (which would have to piece together suitably), given as (say) C^∞ real hypersurfaces in each coordinate patch. In the $\bar{\partial}$-method, the operator H would be required to attain C^∞-smooth boundary values at the boundary of the given $2(n+1)$-real-dimensional C^∞ manifold-with-boundary.

Both methods would provide the boundary $\mathcal{V} = \partial\mathcal{D}$, of the complex manifold $\mathcal{D}$ with a CR-structure, but with the Čech-method, $\mathcal{D}$ would always be locally extendible to the other side of $\mathcal{V}$ (simply by a slight extension of the relevant coordinate patch) whereas with the $\bar{\partial}$-method we would have no guarantee that this can be done. Indeed, the existence of locally nonextendible holomorphic H^1's at (a realizable) $\mathcal{V}$, when there is just one degree of pseudoconcavity, suggests also the existence of locally nonextendible complex manifolds-with-boundary with just one degree of pseudoconcavity. Moreover, with a Levi form of, say, signature $(+-)$, like that of a perturbed version $\mathcal{V} = \mathbb{P}\mathcal{T}_0$ of $\mathbb{P}\mathbb{T}_0$, as described in §3, we may expect that the CR-manifold $\mathcal{V}$ can sometimes be realizable as a boundary *neither* on one side *nor* on the other.

One approach to the construction of such nonrealizable CR-structures is to produce $\mathbb{P}\mathcal{T}_0$, as in §3, where the space-time $\mathcal{M}$ and hypersurface $\mathcal{U}$ are suitably nonanalytic. However the mere nonanalyticity of the resulting CR-structure is not sufficient to ensure its nonrealizability. The reason is, of course, that the real hypersurface defining a realizable CR-manifold need not itself be analytic. Thus, LeBrun (1980) was forced into some more subtle (nonlocal) considerations in order to prove his nonrealizability result for this case.

The related but slightly different earlier suggestion due to Sparling was to use, instead, the Ward (1979) "twisted photon" construction. According to this construction (a special case of Ward's (1977) twistor method of generating self-dual Yang-Mills fields), self-dual Maxwell fields are represented by holomorphic line bundles over the appropriate regions of $\mathbb{P}\mathbb{T}$, rather than by elements $\Psi \in H^1(\mathcal{L}, \mathcal{O}(0))$, as in (4.7). The bundles are constructed using the multiplicative $\mathcal{O}_*$ (zeros excluded), rather than the additive $\mathcal{O}(0)$, the relation between the two being achieved via the familiar exact sequence

$$0 \to \mathbb{Z} \xrightarrow{\times 2\pi i} \mathcal{O}(0) \xrightarrow{\exp} \mathcal{O}_* \to 0. \tag{4.12}$$

Now, real Maxwell fields on (open) regions $\mathcal{R}$ of real Minkowski space M are equivalent to self-dual Maxwell fields on $\mathcal{R}$ (the relevant mutually inverse maps

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

being "take the self-dual part" and "take twice the real part"). To express such fields we can thus either use elements of

$$\hat{H}^1(\mathcal{K}, \mathcal{O}(0)) \tag{4.13}$$

where $\mathcal{K}$ is the region of $\mathbf{PT}_0$ swept out by the projective lines representing the points of $\mathcal{R}$ (as in (4.9)), or else use the corresponding line bundles over $\mathcal{K}$. Such real Maxwell fields certainly need not be analytic, and may even be hyperfunctional. Let us take the field to be C^∞ but nowhere analytic. It may then be represented by a C^∞ CR-line bundle over $\mathcal{K}$ in accordance with the exponentiation procedure of (4.12) as applied to the element of (4.13) representing the field. This bundle is not extendible to a holomorphic line bundle over some open neighbourhood of $\mathcal{K}$ in $\mathbf{PT}$, for if it were, the self-dual Maxwell field on the corresponding open region of $\mathbf{C}M$ that it represents would restrict down to (the self-dual part of) the nonanalytic field that we started with, which is a contradiction since Maxwell fields on open regions of $\mathbf{C}M$ are necessarily holomorphic (and therefore real-analytic). By choosing our original self-dual field to be either of positive or negative frequency, we can ensure *one*-sided extendibility of the resulting bundle into $\mathbf{PT}_-$ or $\mathbf{PT}_+$, respectively (cf. Wells (1981), Bailey, Ehrenpreis and Wells (1982)).

By a slight extension of Sparling's above line of argument, one can translate the above statements concerning inextendible or one-side-extendible CR-line bundles over $\mathcal{K}$ to nonrealizable or one-side-realizable CR-manifolds. The CR-line bundle over $\mathcal{K}$ is itself a CR-manifold with $(+0-)$ Levi signature (and the one direction of holomorphic flatness defines the fibre direction). Now the structure that a holomorphic line bundle $\mathcal{B}$ possesses, over and above that of its complex-manifold structure, can be represented in terms of the holomorphic vector field

$$\lambda \frac{\partial}{\partial \lambda} \tag{4.14}$$

on $\mathcal{B}$, where λ is a fibre coordinate on each fibre. Likewise, the bundle structure of the above CR-line bundle over $\mathcal{K}$—which we now call $\mathcal{V}$—is also characterized by (4.14), which is now a CR-vector field on $\mathcal{V}$. Suppose that $\mathcal{V}$ forms a boundary to a complex manifold-with-boundary $\mathcal{D}$. It follows from the Lewy extension property (and the $(+0-)$ Levi signature) that the vector field (4.14) extends locally into $\mathcal{D}$ and defines $\mathcal{D}$ (locally near $\mathcal{V}$) as a holomorphic line bundle over some open neighbourhood in $\mathbf{PT}_+$ or $\mathbf{PT}_-$ of $\mathcal{K}$. (Though the Lewy extension property refers, in the first instance, only to CR-functions and not to vector fields, the extension to CR-vector fields is easy to achieve by various means, e.g. take components in local CR-coordinates.) It follows from the above arguments (essentially) that our inextendible, or one-side-extendible CR-line bundles are neither-side-realizable or one-side-realizable CR-manifolds, respectively. In the latter case, we have examples of Hill's complex manifolds-with-boundary of the $\bar{\partial}$-type which are not realizable by the $\check{C}$-method, since they are not extendible beyond their C^∞ smooth boundaries.

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

The above discussion has been given partly in outline. It depends to some extent upon the details of twistor geometry and the twistor constructions of physical fields. Not all of this is necessary, however, for the construction of neither-side-realizable and one-side-realizable CR-manifolds. A simplified example due to C. D. Hill, G. A. J. Sparling and the author can be given, where $\mathcal{V} = \partial\mathcal{D}$, with real dimension 5 and Levi signature $(0\ -)$, where the complex 3-manifold-with-boundary $\mathcal{D}$ cannot be extended as a complex manifold beyond its C^∞ boundary $\mathcal{V}$. This example has an advantage over what has been presented above in that the extendibility emerges as manifestly a *local* rather than a global obstruction. (The relation between cohomology classes on portions of **PT** and physical fields depends upon a nonlocal correspondence.) A fully detailed argument will be presented elsewhere, but the gist of the construction is as follows.

Let $\mathcal{S}$ be the unit 3-sphere in $\mathbf{C}^2$, and define $\mathcal{K}$ and $\mathcal{E}$ to be the intersections with $\mathcal{S}$ and with the closed exterior of $\mathcal{S}$, respectively, of the open ball of radius ε (< 2) centered about some point of $\mathcal{S}$. Then $\mathcal{E}$ is a complex 2-manifold-with-boundary, with $\mathcal{K} = \partial\mathcal{E}$ and Levi signature $(-)$. Choose an element $\Phi \in H^1(\mathcal{E} - \mathcal{K}, \mathcal{O})$ which has C^∞ boundary values on $\mathcal{K}$ but which is not locally extendible beyond $\mathcal{K}$. By the aforementioned standard exponentiation process, construct a holomorphic line bundle over $\mathcal{E} - \mathcal{K}$ which joins smoothly to a CR-line bundle over $\mathcal{K}$. These two pieces of line bundle together provide the required complex 3-manifold-with-boundary $\mathcal{D}$ (as its total space) and the portion over $\mathcal{K}$ provides the required C^∞ CR-5-manifold $\mathcal{V}$, of Levi signature $(0\ -)$, beyond which $\mathcal{D}$ is not locally extendible as a complex manifold. For the proof we appeal to the same argument as before, which uses Lewy extension of the vector field (4.14) to show that any extension of $\mathcal{D}$ across $\mathcal{V}$ as a complex manifold is also (locally) an extension of $\mathcal{D}$ as a holomorphic vector bundle. But such extension would imply that Φ extends across $\mathcal{K}$, which is a contradiction.

The Hill philosophy[3] would suggest that a corresponding example should exist in which $\mathcal{D}$ is a 2-complex-dimensional with Levi signature $(-)$ at its C^∞ boundary $\mathcal{V}$. Perhaps the original Nirenberg example bounds on one side and thus provides such a $\mathcal{V}$?

It is remarkable, and perhaps even somewhat ironic that such seemingly esoteric matters as nonrealizable CR-structures and the delicate dividing line between C^ω and C^∞ should have significant connections with physics. But it seems to be so. (The infinite-dimensionality of $H^1(\mathbf{PT}_+, \mathcal{O}(r))$—and hence of physical massless fields—can be attributed to Lewy nonextendibility across $\mathbf{PT}|_{\mathcal{I}}$.) What would Poincaré have made of all this?

I am grateful to T. N. Bailey, M. G. Eastwood, C. D. Hill, C. R. LeBrun and G. A. J. Sparling for many helpful suggestions. I am also grateful to NSF for financial assistance which enabled me to attend the Poincaré Symposium, and for

[3]Recently a large class of nonrealizable CR-manifolds has been presented by Jacobowitz and Trèves (1982). Also Kuranishi (1982) has shown that in dimension $2n + 1 > 7$ and for positive-definite Levi form all CR-manifolds are realizable. All these results are consistent with the Hill philosophy.

 Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

support under contract MCS 79-12938 while the author was at the Institute for Advanced Study, Princeton in March 1980, where some of this research was done.

References

A. Andreotti and C. D. Hill (1972), *E. E. Levi convexity and the Hans Lewy problem. Part I: Reduction to vanishing theorems; Part II: Vanishing theorems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **26**, 325–363; (4) **26**, 747–806.

T. N. Bailey, L. Ehrenpreis and R. O. Wells, Jr. (1982), *Weak solutions of the massless field equations*, Proc. Roy. Soc. London Ser. A (to appear).

S. Bochner (1943), *Analytic and meromorphic continuation by means of Green's formula*, Ann. of Math. (2) **44**, 652–673.

H. Bondi (1965), *Relativity and common sense*, Heinemann, London.

É. Cartan (1923), Ann. Sci. École Norm. Sup. **40**, 325.

_____ (1924), Ann. Sci. École Norm. Sup. **41**, 1.

_____ (1932), *Sur la géométrie pseudo-conforme des hypersurfaces de deux variables complexes*. I, II, Ann. Math. Pura Appl. (4) **11**, 17–90; Ann. Scuola Norm. Sup. Pisa (2) **1**, 333–354.

S. S. Chern and J. K. Moser (1974), *Real hypersurfaces in complex manifolds*, Acta Math. **133**, 219–271.

M. G. Eastwood, R. Penrose and R. O. Wells, Jr. (1981), *Cohomology and massless fields*, Comm. Math. Phys. **78**, 305–351.

K. Friedrichs (1928), *Eine invariante Formulierung des Newtonschen Gravitationsgesetzes und des Grenzüberganges vom Einsteinschen zum Newtonschen Gesetz*, Math. Ann. **98**, 566.

R. P. Geroch (1970), *Domain of dependence*, J. Math. Phys. **11**, 437.

C. D. Hill (1973), *The Cauchy problem for $\bar{\partial}^n$*, Partial Differential Equations, Proc. Sympos. Pure Math., vol. 23, Amer. Math. Soc., Providence, R.I.

C. D. Hill and B. MacKichan (1977), *Hyperfunction cohomology classes and their boundary values*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser. IV **4**, 577–597.

L. Hörmander (1966), *An introduction to complex analysis in several variables*, Van Nostrand-Reinhold, New York.

H. Jacobowitz and F. Trèves (1982), *Non-realizable CR-structures* (to appear).

M. Kashiwara (1979), *Introduction to the theory of hyperfunctions*, Seminar on Micro-local Analysis, Ann. of Math. Studies, no. 93, Princeton Univ. Press, Princeton, N.J.

M. Kuranishi (1982), *Strongly pseudoconvex CR-structures over small balls. Part III: An embedding theorem* (to appear).

C. R. LeBrun (1980), *Spaces of complex geodesics and related structures*, Ph.D. thesis, Oxford Univ.

_____ (1982), *$\mathcal{H}$-space with a cosmological constant*, Proc. Roy. Soc. London Ser. A **380**, 171–185.

J. Leray (1952), *Hyperbolic differential equations*, Institute for Advanced Study.

H. Lewy (1956), *On the local character of a solution of an atypical linear differential equation in three variables and a related theorem for regular functions of two complex variables*, Ann. of Math. (2) **64**, 514–522.

_____ (1957), *An example of a smooth linear partial differential equation without solution*, Ann. of Math. (2) **66**, 155–158.

H. Minkowski (1908), *Space and time*, The Principle of Relativity (H. A. Lorentz, A. Einstein, H. Minkowski and H. Weyl, eds.), Dover, New York, 1923.

A. Newlander and L. Nirenberg (1957), *Complex analytic coordinates in almost complex manifolds*, Ann. of Math. (2) **65**, 391–404.

L. Nirenberg (1973), *Lectures on linear partial differential equations*, CBMS Regional Conf. Ser. in Math., no. 17, Amer. Math. Soc., Providence, R.I.

_____ (1974), *On a question of Hans Lewy*, Russian Math. Surveys **29**, 251–262.

R. Penrose (1959), *The apparent shape of a relativistically moving sphere*, Proc. Cambridge Philos. Soc. **55**, 137–139.

_____ (1968), *Structure of space-time*, Battelle Recontres, 1967 Lectures in Mathematics and Physics (C. M. DeWitt and J. A. Wheeler, eds.), Benjamin, New York.

_____ (1969), *Solutions of the zero rest-mass equations*, J. Math. Phys. **10**, 38–39.

_____ (1975), *Twistor theory, its aims and achievements*, Quantum Gravity, an Oxford Symposium (C. J. Isham, R. Penrose and D. W. Sciama (eds.)), Clarendon Press, Oxford.

Physical space-time and nonrealizable CR-structures. *Bull. Amer. Math. Soc. (N.S.)* **8** (1983), no. 3, 427–48. Also in *The Mathematical Heritage of Henri Poincaré, Proc Symp. Pure Mathematics of the American Mathematical Society*, Vol. 39. Ed. Felix E. Browder, 401–22. American Mathematical Society, 1983

_____ (1977), *The twistor programme*, Rep. Mathematical Phys. **12**, 65–76.

R. Penrose and R. S. Ward (1980), *Twistors for flat and curved space-time*, General Relativity and Gravitation, Vol. II (A. Held (ed.)), Plenum Press, New York and London.

H. Poincaré (1905), *Science and hypothesis*, Walter Scott, London; reprinted, Dover, New York, 1952.

_____ (1906), *Sur la dynamique de l'électron*, Rend. Circ. Mat. Palermo **21**, 129–176.

_____ (1907), *Les functions analytique de deux variables et la représentation conforme*, Rend. Circ. Mat. Palermo **23**, 185–220.

_____ (1954), *Oeuvres de Henri Poincaré*, vol. 9, Gauthier-Villars, Paris.

J. C. Polking and R. O. Wells, Jr. (1976), *Boundary values of Dolbeault cohomology classes and a generalized Bochner-Hartogs theorem*.

M. Sato (1959–60), *Theory of hyperfunctions*, J. Fac. Sci. Univ. Tokyo Sect. I **8**, 139–193, 387–436.

J. L. Synge (1960), *Relativity: the general theory*, North-Holland, Amsterdam.

N. Tanaka (1962), *On the pseudo-conformal geometry of hypersurfaces of the space of n complex variables*, J. Math. Soc. Japan **14**, 397–429.

J. Terrell (1959), Phys. Rev. **116**, 1041.

A. Trautman (1966), *Comparison of Newtonian and relativistic theories of space-time*, Perspectives in Geometry and Relativity (B. Hoffmann (ed.)), Indiana Univ. Press, Bloomington and London.

R. S. Ward (1977), *On self-dual gauge fields*, Phys. Lett. **61A**, 81.

_____ (1979), *The twisted photon: massless fields as bundles*, Advances in Twistor Theory (L. P. Hughston and R. S. Ward (eds.)), Research Notes in Math. 37, Pitman Adv. Publ. Program, San Francisco, London, Melbourne.

R. O. Wells, Jr. (1979), *Complex manifolds and mathematical physics*, Bull. Amer. Math. Soc. (N.S.) **1**, 296–336.

_____ (1981), *Hyperfunction solutions of the zero-rest-mass field equations*, Comm. Math. Phys. **78**, 567–600.

MATHEMATICAL INSTITUTE, OXFORD, GREAT BRITAIN

146
QUASI-LOCAL MASS AND ANGULAR MOMENTUM

10th International Conference on General Relativity and Gravitation, Vol. 1. Eds B. Bertotti, F. De Felice, and A. Pascolini, 607–8. Consiglio Nazionale Delle Ricerche, Rome, 1983

306 Quasi-local mass and angular momentum. *10th International Conference on General Relativity and Gravitation*, Vol. 1. Eds B. Bertotti, F. De Felice, and A. Pascolini, 607–8. 1983

Quasi-Local Mass and Angular Momentum

Roger Penrose
Mathematical Institute
University of Oxford

Given a spacelike 2-surface $\mathcal{S}$, with S^2-topology, a definition is provided, in a general curved space-time, for the total rest-mass $m(\mathcal{S})$ and the ten components of energy-momentum and angular momentum $A_{\alpha\beta}(\mathcal{S})$ surrounded by $\mathcal{S}$. This definition agrees with the Bondi-Sachs mass-momentum at $\mathscr{I}^+$, with the ADM mass at spacelike infinity and with the expected answer for all weak fields with sources (unlike, say, the Komar integral, or previous definitions of angular momentum at $\mathscr{I}^+$, which do not always do so). Tod has computed $m(\mathcal{S})$ in numerous circumstances, e.g. for $\mathcal{S}$ lying in a hypersurface of time-symmetry, and finds that it yields a physically acceptable answer for the mass in all circumstances where his methods can be applied.

The procedure for computing $A_{\alpha\beta}(\mathcal{S})$ and $m(\mathcal{S})$ is first to construct the 2-surface twistor space $\mathbb{T}^\alpha(\mathcal{S})$ as the solution space of the coupled differential equations on $\mathcal{S}$:

$$\eth' \omega^0 = \sigma' \omega^1, \quad \eth \omega^1 = \sigma \omega^0$$

(using standard G.H.P. notation, with spin-frame naturally associated with $\mathcal{S}$). Then, using two elements of $\mathbb{T}^\alpha(\mathcal{S})$ (or, more generally, one element of the symmetric tensor product $\mathbb{T}^{(\alpha\beta)}(\mathcal{S})$), we can reduce the "spin-2" curvature tensor R_{abcd} to a "spin-1" quantity resembling a Maxwell field. The charge integrals for this resulting Maxwell field provide us with the desired element $A_{\alpha\beta}(\mathcal{S}) \in \mathbb{T}_{(\alpha\beta)}(\mathcal{S})$. With a suitable definition of Hermitian (++--)-signature norm on $\mathbb{T}^\alpha(\mathcal{S})$, the quantity

$$m(\mathscr{S}) = (-\tfrac{1}{2}A_{\alpha\beta}(\mathscr{S})\bar{A}^{\alpha\beta}(\mathscr{S}))^{\frac{1}{2}}$$

can be defined.

The major outstanding problems with this approach seem to be the following:

(1) Find a more convincing norm on $\mathbb{T}^{\alpha}(\mathscr{S})$ than the one that exists at present. (Tod's successful examples and the mass defined at infinity are insensitive to this problem.)

(2) Show that $A_{\alpha\beta}(\mathscr{S})$ satisfies a Hermiticity property which reduces the ten <u>complex</u> components of $A_{\alpha\beta}(\mathscr{S})$ to ten <u>real</u> ones. (This problem is intimately related to (1) above.)

(3) Find a (Witten-type?) positivity theorem for $m(\mathscr{S})$, or other physically desirable inequalities.

(4) Find a flux law for angular momentum at $\mathscr{I}^{+}$ (and elsewhere).

(5) Extend Tod's remarkable result that $m(\mathscr{S})$, for the Schwarzschild solution with $\mathscr{S}$ in t = const. (or r = const., or retarded time = const.), depends only on how $\mathscr{S}$ links the source (giving zero if $\mathscr{S}$ does not link the source and the Schwarzschild mass if it links it just once).

(6) Decide whether "exceptional" circumstances can exist in which $\mathbb{T}^{\alpha}(\mathscr{S})$ is more than four-dimensional.

<u>References</u>

Geroch, R., Held, A. & Penrose, R. (1973) J. Math. Phys. <u>14</u>, 874.

Penrose, R. (1982) Proc. Roy. Soc. (Lond.) A <u>381</u>, 53.

Tod, K.P. (1983) Proc. Roy. Soc. (Lond.) A (to appear).

Penrose, R. & Rindler, W. (1983) <u>Spinors and Space-Time, Vol.2: Spinor and Twistor Methods in Space-Time Geometry</u>, (Cambridge University Press).

147
RELATIVE COHOMOLOGY, GOOGLIES AND DEFORMATIONS OF $\mathbb{I}$

Twistor Newsletter **17** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

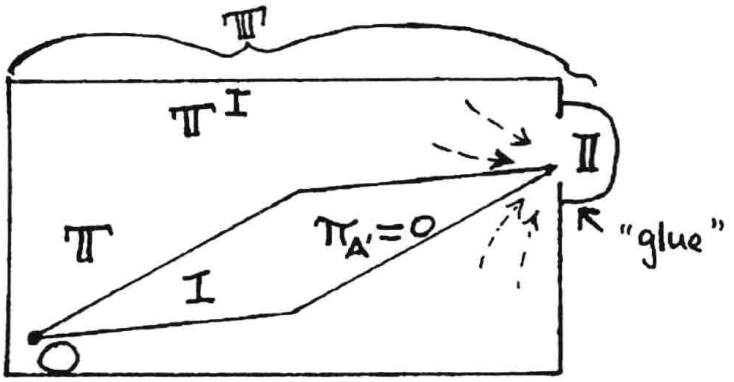
 Relative cohomology, googlies and deformations of $\mathbb{I}$. *Twistor Newsletter* **17** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.12 **Relative Cohomology, Googlies and Deformations of $\mathbb{I}$** *by R. Penrose* (TN 17, January 1984)

In §III.4.11 I suggested that a possible home for the gravitational 'googly' information might lie in the differing (allowable) ways that $\mathbb{I}$ (= the non-projective blown-up line I) can be attached ('glued') to $\mathbb{T}$ (or to $\mathbb{T}^I = \mathbb{T} \setminus I$ — it makes little difference). (Note: $I = \{(\omega^A, \pi_{A'}) | \pi_{A'} = 0\}$.) I assume that the twistor space $\mathbb{T}$ is flat, i.e., that the associated space-time is *self*-dual, with the usual handedness conventions ($\Psi_{ABCD} = 0$, $\tilde{\Psi}_{A'B'C'D'} \neq 0$).

Recall that $\mathbb{I}$ is given by the finite limits of

$$(\omega^A \pi_{B'}, \pi_{A'} \pi_{B'})$$



as $\pi_{A'} \to O, \omega^A \to \infty$. We expect that the allowable deformations of the gluing of $\mathbb{I}$ to $\mathbb{T}$ will be alterations in the *holomorphic* structure of $\mathbb{T} = \mathbb{T} \cup \mathbb{I}$ (not affecting which points of $\mathbb{I}$ are topological limits of which point sequences in $\mathbb{T}$). An infinitesimal allowable deformation of this gluing would be expected to show up as a relative cohomology group (relative to $\mathbb{I}$) — in fact an $H^2_{\mathbb{I}}$ of holomorphic vector fields in $\mathbb{T}$ which are tangent to $\mathbb{I}$ at $\mathbb{I}$. I shall address the problem of how such an $H^2_{\mathbb{I}}$ might be obtainable from an $H^1(\cdots, \mathcal{O}(-6))$ — which represents a linearized self-dual gravitational field in the now standard way. Consider the exact sequence

$$0 \to \widehat{d\Omega^2} \xrightarrow{i} \widehat{\Omega^3} \xrightarrow{d} \Omega^4 \to 0\,, \tag{1}$$

where $\widehat{\Omega^3}$ denotes the sheaf of holomorphic 3-forms on $\mathbb{T}$ which, on $\mathbb{T}$, are proportional to $\mathscr{D}Z = \frac{1}{6}\widehat{Z dZ \wedge dZ \wedge dZ}$ those which are closed being denoted $\widehat{d\Omega^2}$; and where Ω^4 denotes holomorphic 4-forms on $\mathbb{T}$. Recall that $\pi_{A'}\pi_{B'}\mathscr{D}Z$ is finite on $\mathbb{T}$ and

$$d(\pi_{A'}\pi_{B'} f \mathscr{D}Z) = \pi_{A'}\pi_{B'} d^4Z(\Upsilon + 6)f\,, \tag{2}$$

Υ being the Euler operator and $d^4Z = \frac{1}{24}\widehat{dZ \wedge dZ \wedge dZ \wedge dZ}$. From (1),

$$\cdots \to H^1(\mathcal{Q}, \widehat{\Omega}^3) \to H^1(\mathcal{Q}, \Omega^4) \xrightarrow{\int} H^2(\mathcal{Q}, \widehat{d\Omega^2}) \to \cdots \tag{3}$$

is exact, $\mathcal{Q}$ being some suitable neighbourhood of $\mathbb{I}$ (or perhaps of part of $\mathbb{I}$) in $\mathbb{T}$. Note that the map $\int$ does indeed involve an integration of quantities of the form gd^4Z (think of what the connecting homomorphism actually means) — and for regularity at $\mathbb{I}$ we must have two π's in g; so what is involved is the inverse operation in (2), i.e., '$(\Upsilon + 6)^{-1}$', on quantities of that form for which this operation does not globally exist,

 Relative cohomology, googlies and deformations of $\mathbb{I}$. *Twistor Newsletter* **17** (1984).
Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

i.e., degree -6 functions. For the other homogeneity degrees in $H^1(\mathcal{Q}, \Omega^4)$, these parts come from the *left* in (3), whereas the -6's go to the *right*. (Note: $(\Upsilon + 6)^{-1}$ acting on a -6 function is that function $\times$ a logarithm. The logarithm would need to extend to $\mathbb{I}$ which is why the H^2 comes in instead here.)

Now from the relative cohomology sequence we have

$$\cdots \to H^1(\mathcal{Q} - \mathbb{I}, \widehat{d\Omega^2}) \to H^2_{\mathbb{I}}(\mathcal{Q}, \widehat{d\Omega^2}) \to H^2(\mathcal{Q}, \widehat{d\Omega^2}) \to \cdots$$

Again it is -6 functions that are relevant here, since these provide $\widehat{d\Omega^2}$ finite on $\mathbb{I}$, with zero on the right in (2) for closed-ness. The idea is that the appropriate part of $H^2_{\mathbb{I}}(\mathcal{Q}, \widehat{d\Omega^2})$ should 'be' $H^2(\mathcal{Q}, \widehat{d\Omega^2})$ which by (3) is provided by -6 homogeneity twistor functions (H^1's). (It thus appears that the required infinitesimal deformations can be provided by -6 functions $\times$ logarithms.) Note that the 3-forms in $\widehat{\Omega^3}$ (and in particular $\widehat{d\Omega^2}$) 'point' in the direction of Υ, which is tangent to $\mathbb{I}$ at $\mathbb{I}$, as required. (We can divide a 3-form by a 4-form — canonically d^4Z in $\mathbb{T}$ — to get a vector field.)

All this clearly needs much sharpening up. Work is in progress (see §III.4.1). The required deformed gluings of $\mathbb{I}$ to $\mathbb{T}$ are supposed to determine which are the correct googly maps by a regularity condition at $\mathbb{I}$. A possible line of approach for obtaining insights is to examine how twistor functions (perhaps of the dual Helfer variety, §III.4.8) can be used to yield test fields on the googly space-time via contour integrals involving these googly maps. There appears to be some interesting topology.

Thanks to T.N. Bailey especially.

148
DONALDSON'S MODULI SPACE: A 'MODEL' FOR QUANTUM GRAVITY?

Quantum theory of gravity (1984), 295–8, Hilger, Bristol

 Donaldson's moduli space: A 'model' for quantum gravity? *Quantum theory of gravity*, 295–8, 1984

Donaldson's Moduli Space: a 'Model' for Quantum Gravity?

ROGER PENROSE

I fear that I must apologize to Bryce DeWitt for the thoughts which follow, since though dedicated to him, they are of the nature merely of vaguest speculation rather than being a contribution to the thorough-going, well founded type of approach to quantum gravity of which he has been a pioneer and major exponent. These thoughts do, however, owe to him the helpful suggestion that insights may be gained by studying the simpler types of non-linearity which occur in Yang–Mills theory as a preliminary to tackling those more complicated and elusive ones of general relativity (cf DeWitt 1964).

For various reasons, which I do not propose to go into here (cf Penrose 1976, 1979, 1981, 1982, also Karolyhazy 1966, Mielnick 1974, Komar 1969, Kibble 1981), I have found myself driven to take a decidedly unorthodox viewpoint with regard to the features I would expect (or hope) to be present in the 'true' quantum gravity theory. Among such features are:

(a) time-asymmetry (including effective CPT-violation),
(b) a non-linear modified quantum theory,
(c) an objective theory to take over the role of subjective 'wavefunction reduction',
(d) space time singularities not to be replaced by non-singular 'bounces'; instead, insofar as a space–time manifold remains a good approximation, the Weyl curvature should approach zero at initial singular boundary points.

It seems highly unlikely that any too 'orthodox' approach to quantum gravity could produce anything much resembling (a), . . . , (d)—which, I suspect, would be reason enough for most people simply to reject these desiderata and go on their own merry ways. Nevertheless, I strongly feel that the evidence from other sources (primarily the existence of a second law of thermodynamics, coupled with cosmological evidence constraining the structure of the big bang singularity) is sufficiently persuasive to indicate that something of the nature of (a), . . . , (d) ought to be striven for.

This suggests to me that while 'orthodox' quantization may supply necessary clues, we should look beyond these standard procedures for the eventual formulation of a quantum gravity theory. Rather than a theory obtainable by 'correct' application of standard quantization to standard general relativity, I would anticipate something more in the nature of a 'grand synthesis', which would be a theory having standard quantum theory and standard general relativity as two different limits (i.e. probably the limits of small energies and of large dimensions). Since this means we are concerned with a theory whose structure is different both from that of a Hilbert space and that of a (pseudo-) Riemannian manifold, any insights into possible mathematical objects which might, even if only very superficially, possess the right kind of limiting behaviors would be something worthy of serious examination. It is in this spirit that I put forward the following highly tentative 'model'. It is not at all to be thought of as a proposal for a quantum gravity theory (which it certainly is not), but as something possibly analogous, in some limited but possibly suggestive ways, to such a theory.

The 'model' I am bringing forward is a construction (due to Donaldson 1983) which has, over the past year, caused a certain amount of excitement in pure mathematical circles, since it has led to the proof of an outstanding (and not altogether anticipated) result concerning the topology of compact differentiable 4-manifolds. When combined with other results one obtains, as a corollary, the remarkable fact that $\mathbb{R}^4$ possesses more than one distinct differentiable structure—a property not shared by $\mathbb{R}^n$ for any other value of n (cf Donaldson 1983, and Atiyah 1983 for an overall view).

Let $\mathscr{M}$ be a compact differentiable 4-manifold which is topologically 'positive' in the sense that the intersection matrix of its 2-cycles is positive-definite. Supply it with a C^∞ Riemannian (positive-definite) metric g. We look for solutions of SU(2)–Yang–Mills theory on $\mathscr{M}$ which are self-dual and have instanton number one. Let $\mathscr{D}$ be the moduli-space of these solutions (i.e. the points of $\mathscr{D}$ represent solutions under gauge equivalence; so distinct points of $\mathscr{D}$ represent Yang–Mills instantons not related to one another by a gauge transformation). Now $\mathscr{D}$ may possess singular points, of one or other of two types A or B, but apart from that it is a smooth 5-manifold-with-boundary. The B-singularities arise only for 'special' choices of the metric g. If we are concerned only with topological (or differential-topological) questions concerning $\mathscr{M}$, then the particular choice of metric g that has been made is of no consequence. We need only ensure that g is suitably 'generic' and all B-singularities will disappear.

The A-singularities remain, however. These consist of a certain number, N, of isolated points within $\mathscr{D}$, where N is the second Betti number of $\mathscr{M}$. Locally, each A-point is the vertex of a cone whose cross-section is $\mathbb{CP}^2$ (complex projective two-space). The boundary $\partial\mathscr{D}$ of $\mathscr{D}$ turns out to be diffeomorphic to $\mathscr{M}$. We may envisage a truncated manifold $\hat{\mathscr{D}}$, obtained from $\mathscr{D}$ by 'cutting off' all the A-points, so that $\hat{\mathscr{D}}$ is, indeed, a smooth manifold-with-boundary, the boundary being diffeomorphic to a disjoint union:

$$\underbrace{\mathbb{CP}^2 \cup \ldots \cup \mathbb{CP}^2}_{N \text{ copies}} \cup \mathscr{M}$$

CHAPTER 148

 Donaldson's moduli space: A 'model' for quantum gravity? *Quantum theory of gravity*, 295–8, 1984

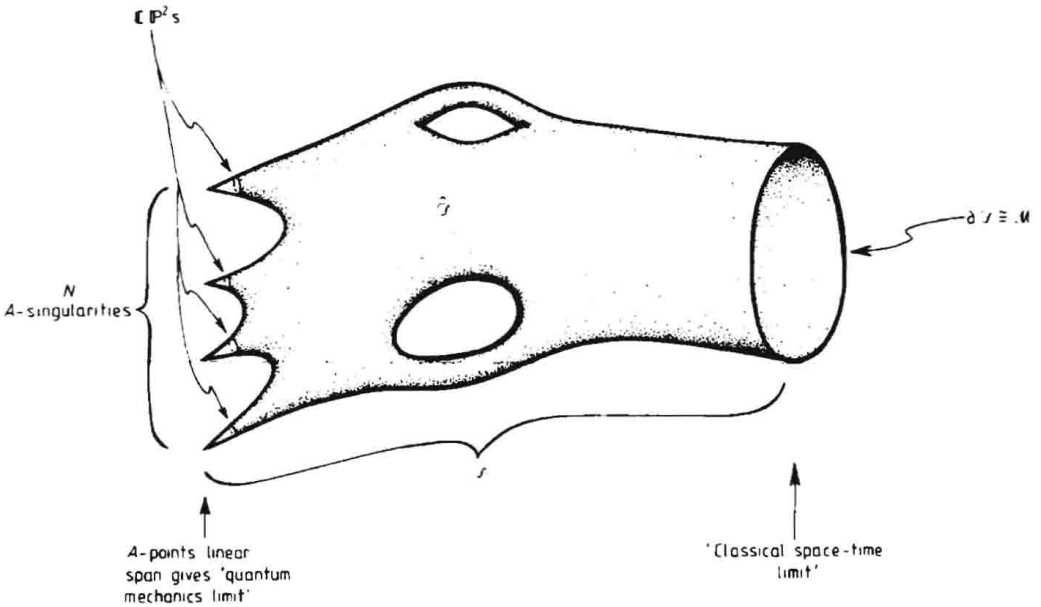


Figure 1 Donaldson's moduli space for SU(2)-instantons, taken as a tentative 'model' for a quantum gravity theory.

This supplies, in a concrete way, a *cobordism* between $\mathcal{M}$ and n copies of $\mathbb{CP}^2$ (see figure 1). Moreover, each $\mathbb{CP}^2$ turns out to count with the same orientation, and from this the required topological theorem can be derived (that the intersection matrix of two-cycles is diagonal).

The significance (if any) for a quantum gravity program lies in the interpretation of the $\mathcal{A}$-points and the identity of $\mathcal{M}$ with $\hat{\partial}\mathcal{Y}$. The $\mathcal{A}$-points arise from those particular SU(2)-instantons which 'degenerate' to the *linear* Maxwell case (with constant 'Yang–Mills direction'). The relation to the second Betti number comes from de Rham–Hodge theory. The boundary $\hat{\partial}\mathcal{Y}$ arises from those limiting instantons which are concentrated at points of $\mathcal{M}$. I would think of the $\mathcal{A}$-points as the analogs of 'physically realized quantum states' or a 'natural basis' for the quantum Hilbert space, where the full Hilbert space is given as their linear span. This Hilbert space corresponds, in some sense, to the linear limit of the Yang–Mills theory. On the other hand the boundary $\hat{\partial}\mathcal{Y}$ is viewed as corresponding to the classical space time (general relativity) limit of the theory, with individual instantons concentrated at points.

The idea would be that the 'correct' quantum gravity theory would in some sense correspond to the *interior* points of $\hat{\mathcal{Y}}$, so that neither the purely quantum-mechanical nor the classical general-relativistic picture would ever be strictly accurate ($\mathcal{A}$-singularities and boundary points being removed) but could be approached arbitrarily closely in a limiting sense.

One can raise many objections to this as a serious model for a quantum gravity picture (and only (b) of my list is in any way incorporated), but it might be worth thinking about further.

ACKNOWLEDGMENT

I am grateful to Michael Atiyah for a beautifully clear exposition of the Donaldson scheme.

REFERENCES

Atiyah M F 1983 Report to 1982 Durham conference
DeWitt B S 1964 in *Relativity, Groups and Topology: The* 1963 *Les Houches Lectures* ed B DeWitt and C DeWitt (New York: Gordon and Breach)
Donaldson S K 1983 *Bull. Am. Math. Soc.* **8** 81
Karolyhazy F 1966 *Nuovo Cimento* A **42** 390
Komar R 1969 *Int. J. Theor. Phys.* **2** 157
Kibble T W 1981 in *Quantum Gravity 2* ed C J Isham, R Penrose and D W Sciama (Oxford: Oxford University Press)
Mielnick B 1974 *Commun. Math. Phys.* **37** 221
Penrose R 1976 *Gen. Relativ. Gravity* **7** 31
—— 1979 in *General Relativity, An Einstein Centenary Survey* ed S W Hawking and W Israel (Cambridge: Cambridge University Press)
—— 1981 in *Quantum Gravity 2* ed C J Isham, R Penrose and D W Sciama (Oxford: Oxford University Press)
—— 1982 in *Quantum Structure of Space and Time* ed M J Duff and C J Isham (Cambridge: Cambridge University Press)

149
MASS IN GENERAL RELATIVITY

Global Riemannian geometry (1983). Eds T. J. Wilmore and N. J. Hitchin, 203–13. Ellis Horwood Ser. Math. Appl., Chichester, 1984

 Mass in general relativity. *Global Riemannian geometry* (1983). Eds T. J. Wilmore and N. J. Hitchin, 203–13, 1984

Mass in General Relativity

R. Penrose
The Mathematical Institute, Oxford University

1. SOURCE INTEGRALS

In Maxwell theory, there are simple space-time integral expressions for the electric charge of a system. This is because the Maxwell equations take the simple exterior derivative form

$$\mathrm{d}\boldsymbol{F} = 0, \qquad \mathrm{d}\boldsymbol{F}^* = 8\pi\boldsymbol{J}^*, \tag{1}$$

$\boldsymbol{F}$ being an electromagnetic 2-form, $\boldsymbol{F}^*$ its dual and $\boldsymbol{J}^*$ the 3-form dual to the charge–current vector, satisfying

$$\mathrm{d}\boldsymbol{J}^* = 0. \tag{2}$$

The total net charge Q intercepted by a hypersurface $\mathscr{H}$ in the space-time is given by

$$Q = \int_{\mathscr{H}} \boldsymbol{J}^* \tag{3}$$

which, provided that $\mathscr{H}$ is compact with boundary

$$\mathscr{S} = \partial\mathscr{H}, \tag{4}$$

can be re-expressed entirely in terms of the Maxwell field at S, as

$$Q = \frac{1}{8\pi} \oint_{\mathscr{S}} \boldsymbol{F}^* \tag{5}$$

(by (1)). This is the total electric charge surrounded, or 'linked', by the closed 2-surface $\mathscr{S}$, and we do not need to know the details of this charge–current distribution inside $\mathscr{S}$. Similarly the vanishing of magnetic charge surrounded by any such 2-surface is expressed by

$$\oint_{\mathscr{S}} \boldsymbol{F} = 0. \tag{6}$$

The expressions (5) and (6) hold in a general curved space-time $\mathscr{M}$ just as

well as in Minkowski space $\mathbb{M}$, where if there are topologically non-trivial 2-cycles, (6) would normally be postulated for these cycles (since it would not now follow from (1)) or else derived from the existence of a potential.

Charge conservation shows up in two ways, with the expression (5). In part, it is implicit in the very existence of such an expression, since the total charge depends only on the values of $\boldsymbol{F}$ at $\mathscr{S}$ and is the same whichever hypersurface is chosen to span $\mathscr{S}$. The total charge intercepted by an 'earlier' such hypersurface is the same as that intercepted by a 'later' one. But in addition to this we have the fact that if the surface $\mathscr{S}$ is moved continuously through *charge-free* space, then the value of Q, as given by (5), does not change.

In gravitation theory, the most direct analogue of charge is mass. So we may ask whether there is any direct analogue of (3), or of (5), giving the total mass of a system surrounded by some closed 2-surface $\mathscr{S}$. However, there are good physical reasons why we must expect some severe complications to this elegant picture. In the first place, mass is not simply an additive scalar quantity like electric charge. For if by 'mass' we mean 'rest mass', it is not additive (the rest mass of the sum is not the sum of the rest masses, even in special relativity, but involves the relative velocities), while if we mean 'inertial mass' or 'energy', it is merely one component of the energy–momentum 4-vector. In general relativity we have difficulty defining a concept of '4-vector relative to the surface $\mathscr{S}$' – let alone globally.

Moreover, we should really be concerned not just with the four components of energy–momentum but with the 10 of energy–momentum–angular momentum – which I shall refer to collectively as the *kinematics*† of the system. Each of these 10 components is associated with a generator of the Poincaré group (the group of isometries of $\mathbb{M}$), energy and momentum being associated with translations and angular momentum with (Lorentz) rotations. These generators are the Killing vectors of $\mathbb{M}$. If $\boldsymbol{T}$ is a (symmetric, divergence-free) energy–momentum tensor in $\mathbb{M}$, then its contraction $\boldsymbol{L}$ with such a Killing vector $\boldsymbol{K}$ is a covector which is *conserved* in the same sense as was $\boldsymbol{J}$ in (2):

$$T_{ab} = T_{ba}, \qquad \nabla^a T_{ab} = 0 \tag{7}$$

and

$$\nabla_{(a} K_{b)} = 0 \tag{8}$$

together imply

†In the older literature, the term 'kinematics' normally refers to concepts such as position, velocity and angular velocity, but in more recent years, as phase space has grown in importance at the expense of configuration space, the concepts of energy, momentum and angular momentum have also been deemed to be part of kinematics, the term 'dynamics' applying only when a Hamiltonian, Lagrangian, etc., has been specified.

 Mass in general relativity. *Global Riemannian geometry* (1983). Eds T. J. Wilmore and N. J. Hitchin, 203–13, 1984

$$\nabla^a L_a = 0, \tag{9}$$

where

$$L_a = T_{ab}K^b, \tag{10}$$

i.e.

$$\mathrm{d}\boldsymbol{L}^* = 0 \tag{11}$$

with

$$L^*_{abc} = e_{abcd}L^d. \tag{12}$$

(Here ∇ is covariant derivative, round brackets around indices denote symmetric part, $\boldsymbol{e}$ is the Levi–Civita alternating tensor.) $\boldsymbol{L}$ is the conserved current describing that particular component of the kinematics which is associated with $\boldsymbol{K}$, and its total value P intercepted by the hypersurface $\mathscr{H}$ is given, as in (3), by

$$P = \int_{\mathscr{H}} \boldsymbol{L}^*. \tag{13}$$

2. GENERAL RELATIVITY

We seek an analogue of (5) which gives meaning to P also in the curved space-times of general relativity, where we do not normally have solutions of (8) (although (7) will still hold). We cannot now expect to have a conservation law in the strong second sense that the value of P be unaltered when the surface $\mathscr{S}$ is moved through matter-free regions – although there is nothing against conservation in the first sense, that P be independent of any particular choice of $\mathscr{H}$ spanning $\mathscr{S}$. The physical reason for this is that the gravitational field in empty space must itself be able to carry energy (and hence momentum and angular momentum) in the form of gravitational waves. (Indeed this is now apparently an observational fact witnessed in the speeding up of the binary pulsar (cf. [21]). Related to this is the more primitive consideration that gravitational potential energy contributes (negatively) to the total mass of a system: the total mass of two gravitating bodies momentarily at rest with respect to one another must be less than the sum of the masses of the bodies individually by an (in practice very small) amount equal to the gravitational potential energy of interaction (in units where the velocity of light $c = 1$). In each case the empty-space contribution to the total energy shows up non-locally. No flux quantity like $\boldsymbol{L}$ can arise which locally describes the gravitational contribution to the total energy. A precise expression has been known for some time which measures the flux *at infinity* of energy due to gravitational waves [3,18] (see also [11,13]), but this expression has the property that it cannot be determined locally (even at infinity) in terms of the space-time geometry, since it can be non-zero in locally *flat* space-time (cf. [12,17]).

Moreover, the very form of Einstein's equations (with gravitational constant G),

$$R_{ab} - \tfrac{1}{2}Rg_{ab} = -8\pi G T_{ab}, \tag{14}$$

in which the T_{ab} on the right is supposed to describe *all* the local energy, shows that any additional contribution to the total energy arising from the vacuum gravitational field (where $T_{ab} = 0$) can only arise non-locally. Indeed it is the fact that T_{ab}, subject to (7), does not *in itself* give rise to a conserved total energy–momentum flux (that is, in the absence of Killing vectors $\boldsymbol{K}$) which allows a gravitational contribution to the total mass–energy to arise.

In the early days of general relativity, gravitational energy was described by means of a 'pseudo-tensor', which was a coordinate-dependent object that, when added to T_{ab}, produced a quantity whose coordinate-divergence vanished, in place of the covariant equation (7), so that an integral conservation law for gravitational + matter energy could be derived [4]. The non-locality of gravitational energy then showed up in the 'pseudo-tensor' character of the quantity describing it; for, at any one space-time point, this measure of gravitational energy density could be reduced to zero by suitable choice of coordinates. Only the *total* energy, obtained by integrating these non-tensorial quantities all the way to infinity, was considered to have physical meaning.

However, the extreme coordinate dependence of these quantities was always a severe obstacle to finding a suitable (or believable) interpretation. Even when taken to infinity, coordinate dependence remained, so to obtain meaningful answers for the total energy of a system it was necessary to restrict the coordinates to be of a special asymptotically Cartesian type. This was never a particularly satisfactory procedure and various improvements were obtained, leading to some much better-defined expressions which involved the space-time geometry only at infinity, notably the Bondi–Sachs mass [2,3,18], defined at null infinity, and the ADM mass [1] (see also[10]), defined at space-like infinity. Of these, the Bondi–Sachs mass had the distinct advantage that, for the first time, it was possible to assign a precise measure to the mass and momentum carried away from a system by gravitational waves, a striking feature of this measure being the manifest positive-definiteness of the mass of the waves.

However, no general measure of mass at a *quasi-local* level (the mass surrounded by some arbitrary closed 2-surface $\mathcal{S}$) was obtainable by these methods. But in the particular situation of a space-time which is *static* (i.e. possessing a time-like Killing vector), a different procedure due to Komar [9] led to a quasi-local definition of mass which appears to be satisfactory provided that $\mathcal{S}$ lies entirely in a matter-free region. The Komar mass appears to behave somewhat anomalously, however, when $\mathcal{S}$ passes

 Mass in general relativity. *Global Riemannian geometry* (1983). Eds T. J. Wilmore and N. J. Hitchin, 203–13, 1984

through a matter region (see for example, [22]), presumably due to the fact that the matter contributes to the Komar expression through the Ricci tensor directly rather through its trace-reverse, as would be demanded by (14). Moreover, the Komar expression cannot meaningfully be applied quasi-locally in the absence of a Killing vector. On the other hand, this expression has been successfully applied at *infinity*, where appropriately asymptotic Killing vectors exist, describing asymptotic symmetries of an asymptotically flat space-time $\mathcal{M}$, to re-derive the Bondi–Sachs mass (see [20]).

Another expression of some interest in this connection is one due to Hawking [7]. This provides a quasi-local definition of 'mass' surrounded by a space-like 2-surface $\mathcal{S}$ (with S^2 topology) which reduces to the correct Bondi–Sachs expression when applied at null infinity. However, at the quasi-local level, the Hawking expression cannot be expected to give reasonable answers for the *physical* mass, except for very special ('sufficiently round') surfaces $\mathcal{S}$. In $\mathbb{M}$, for example, this 'Hawking mass' is non-zero in general, though the physical value of the mass should be zero. No appropriate restrictions on $\mathcal{S}$ have been found, as yet, which ensure that the Hawking expression gives the physically desired result (but cf. [8]).

3. A NEW APPROACH

Recently, a suggestion for a quasi-local mass definition has been put forward [15] which avoids these difficulties. Again, agreement with the Bondi–Sachs expression is obtained (and also, as has been recently shown by W. T. Shaw, with the ADM expression). But no Killing vector is needed for the space-time $\mathcal{M}$, and the result for $\mathcal{S}$ in $\mathbb{M}$ is always zero, as it should be. Furthermore, the expression provides not just a quasi-local definition of mass, but of the full kinematics (in the sense described earlier) surrounded by $\mathcal{S}$. The definition of angular momentum at null infinity that is thereby provided appears to be a distinct improvement over the ones which had been suggested earlier (cf. [6,23]) in that agreement is always obtained with the answers known to be correct for the weak-field limit of general relativity, whereas this had not been true of any earlier definition, in the case of an arbitrary surface $\mathcal{S}$ at null infinity. In fact, full agreement with the weak-field kinematics is now achieved also for finite $\mathcal{S}$.

This new definition of quasi-local mass has been examined in detail by Tod, and he finds it to give physically reasonable answers for all cases in which he has been able to compute it [22]. In particular, if $\mathcal{S}$ is constrained to lie in a constant-time hypersurface in the Schwarzschild solution, then the Schwarzschild mass is obtained if $\mathcal{S}$ links the source, and zero is obtained if it does not. (It seems natural to conjecture that this holds also if $\mathcal{S}$ is not so constrained.) More generally, if $\mathcal{M}$ is a vacuum space-time

containing a hypersurface Σ of time-symmetry, then the mass turns out to be unchanged under continuous deformation of $\mathcal{S}$ within Σ. If such $\mathcal{M}$ describes, say, a configuration of several black holes (or 'wormholes') momentarily at rest with respect to one another, then the mass value obtained is a function only of the topological relation of $\mathcal{S}$ ($\subset \Sigma$) to these holes. The mass is not additive, however, but gives a physically expected discrepancy consistent with an additional (negative) gravitational potential energy contribution to the total mass when more than one hole is surrounded. Tod also examines cosmological models and finds, in particular, that for a 2-sphere $\mathcal{S}$ of symmetry in a standard Friedmann–Robertson–Walker model, the mass works out to be the same as it would have been for a standard sphere in Euclidean 3-space, of the same area as $\mathcal{S}$, surrounding a fluid ball of the same density as in the model. This implies that for a spatially closed $S^3 \times \mathbb{R}$ model, the mass increases to a maximum as $\mathcal{S}$ is drawn larger and larger, on $\Sigma \cong S^3$, until it becomes an 'equational' surface. Then the mass decreases back to zero as $\mathcal{S}$ shrinks to a point 'around the back of the universe', illustrating the physically anticipated result that the total mass (matter plus potential energy contribution) of a spatially closed universe must be zero. This also illustrates a general property of the new mass definition, that for a spatially closed $\mathcal{M}$, the mass on one side of $\mathcal{S}$ is always equal to the mass on the other!

Certain problems concerning this new definition do still remain, however. One of these lies in the fact that a definition of Hermitian 'norm', that is needed for $\mathcal{S}$ as a whole, arises defined only pointwise on $\mathcal{S}$, and the suggestion that this should be *averaged* over $\mathcal{S}$ is not altogether compelling, since there are other possibilities. In the cases that Tod has been able to treat successfully, this 'pointwise norm' turns out to be constant over $\mathcal{S}$, so that possible ambiguities (e.g. in the mode of averaging) do not arise. A related problem concerns the fact that the 'kinematics', as defined with respect to $\mathcal{S}$, turn out to be described by 10 *complex* quantities, while the standard kinematical freedom involves only 10 real parameters. It is hoped that some Hermiticity property will reduce these 10 complex quantities to 10 real ones, as is the case in the weak-field limit, or for $\mathcal{S}$ at infinity, and in all cases so far explicitly examined. But a general argument is lacking.

4. EXPLICIT PROCEDURE

I shall indicate, in outline only, the procedure for obtaining this quasi-local mass and kinematics. (For more complete discussions see [15–17, 22]. Basically, what one requires is a way of reducing the 'spin-2 quantity', which is the space-time curvature, to a 'spin-1 quantity' resembling a (dual) Maxwell field, to which the formula (5) can then be applied.

Consider first the weak-field limit of general relativity. This may be

described in terms of a tensor field in $\mathbb{M}$ with the symmetries of a Riemann curvature, subject to the 'Bianchi identities' as a field equation, and related to a divergence-free symmetric energy–momentum tensor in $\mathbb{M}$ by the Einstein field equations. Thus we have H_{abcd}, in $\mathbb{M}$, subject to

$$H_{abcd} = H_{[cd][ab]}, \qquad H_{[abc]d} = 0, \tag{15}$$

$$\nabla_{[a}H_{bc]de} = 0 \tag{16}$$

and (cf. 14))

$$H_{abcd}g^{bd} = -8\pi G\{T_{ab} - \tfrac{1}{2}\, g_{ab}T_{cd}g^{cd}\}, \tag{17}$$

where square brackets around indices denote the antisymmetric part and where g_{ab} now refers to the Minkowski metric of $\mathbb{M}$. The equations (7) on T_{ab} hold as before. It may be shown (most easily using spinor methods; cf. [17]) that the (dual) 'Maxwell field'

$$F^*_{ab} = \frac{1}{G}H_{abcd}S^{cd}, \tag{18}$$

where S^{ab} $(= -S^{ba})$ is subject to

$$\nabla^{(a}S^{b)c} - \nabla^{(a}S^{c)b} + g^{a[b}\nabla_d S^{c]d} = 0 \tag{19}$$

and

$$e_{abcd}\nabla^b S^{cd} = 3/2\, K_a \tag{20}$$

for some given Killing vector $\boldsymbol{K}$ for $\mathbb{M}$, satisfies the second Maxwell equation (1):

$$\nabla_{[a}F^*_{bc]} = 8\pi L^*_{abc}, \tag{21}$$

with $\boldsymbol{L}^*$ given by (10) and (12). Thus, if we substitute the $\boldsymbol{F}^*$ given by (18) into the integral (5), we shall obtain the required total kinematic component P surrounded by S, in agreement with the 3-space integral (13) of $\boldsymbol{L}^*$.

In $\mathbb{M}$, (19) has 20 independent solutions, each of which automatically gives rise to a Killing vector in (20). Moreover, for any *given* Killing vector $\boldsymbol{K}$, (20) and (19) together have 10 independent solutions, each of which yields the same quantity P when $\boldsymbol{F}^*$ is integrated over $\mathcal{S}$. However, in a general space-time $\mathcal{M}$ (19) will have only the trivial solution, since it is a highly over-determined equation. From the physical point of view it is reasonable that (19) should not generally have non-trivial solutions, since any such solution would yield an $\boldsymbol{F}^*$ which, in *vacuum*, would satisfy $\mathrm{d}\boldsymbol{F}^* = 0$. Hence a conservation law of the strong second kind would result, for which no contribution to the quantity P could arise from gravitational waves. Since we are not demanding such a strong conservation law but

merely a *definition* of the quantity P, as surrounded by $\mathscr{S}$, it is reasonable to ask merely that (19) hold, in some sense, only for derivatives which act tangentially to $\mathscr{S}$. But this turns out to be a bit too simple-minded, these equations now being *under*-determined. In order to gain the necessary insights into how to proceed, we must turn to a spinor formalism.

I adopt the standard 2-spinor (abstract index) notation (see [14,17]). Setting

$$S^{ab} = \gamma^{AB}\,\varepsilon^{A'B'} + \varepsilon^{AB}\,\bar{\gamma}^{A'B'} \tag{22}$$

(with $\gamma^{AB} = \gamma^{BA}$), we find that (19) reduces to the '2-index symmetric twistor equation'

$$\nabla^{(A}_{A'}\,\gamma^{BC)} = 0, \tag{23}$$

which in $\mathbb{M}$ has 10 independent complex solutions. In $\mathscr{M}$ it is over-determined (2×4 complex equations for three unknowns), whereas the totally tangential equations are only two in number, which is insufficient. The key idea is not to take (23) over as it stands, but to regard the solutions, in $\mathbb{M}$, of (23) as arising as *symmetric tensor products* of solutions of the standard twistor equation

$$\nabla^{(A}_{A'}\,\omega^{B)} = 0. \tag{24}$$

The solutions of (23), in $\mathbb{M}$, are spanned precisely by the symmetrized tensor products of solutions of (24) with themselves. The solutions of (24) in $\mathbb{M}$ constitute a four-dimensional complex vector space $\mathbb{T}$, referred to as the *twistor space* for $\mathbb{M}$, and the solutions of (23) may be regarded as the elements of the 10-dimensional complex vector space $\mathbb{T} \circledS \mathbb{T}$. (Note: $10 = \frac{1}{2} \times 4 \times 5$.)

In $\mathscr{M}$, there are now *two* independent totally tangential complex equations for the *two* components ω^0, ω^1 of ω^A. Using spin-frames locally on $\mathscr{S}$ so that the flagpole directions of the basis spinors o^A, ι^A are the null directions orthogonal to $\mathscr{S}$ – where $\mathscr{S}$ is taken to be necessarily space-like – we find, using the compacted spin-coefficient formalism [5,17], that these two tangential equations can be written

$$\eth'\omega^0 = \sigma'\omega^1, \qquad \eth\omega^1 = \sigma\omega^0. \tag{25}$$

The precise definitions of the quantities in (25) are given in the literature cited, but roughly speaking $\eth$ and $\eth'$ are derivative operators in the holomorphic and antiholomorphic directions within $\mathscr{S}$, as applied to the various line bundles naturally associated with $\mathscr{S}$. The complex quantities σ and σ' describe the *shear*, at $\mathscr{S}$, of the null geodesics orthogonal to $\mathscr{S}$; i.e. they provide the components of the trace-free (conformally invariant) part of the extrinsic curvature of $\mathscr{S}$ in $\mathscr{M}$.

By the Atiyah–Singer index theorem it follows that the coupled

 Mass in general relativity. *Global Riemannian geometry* (1983). Eds T. J. Wilmore and N. J. Hitchin, 203–13, 1984

equations (25) have precisely a four-complex-dimensional space $\mathbb{T}(\mathcal{S})$ of solutions in general (and perhaps always) *provided that the space-like surface $\mathcal{S}$ has S^2-topology*. We assume this to be the case, and find that, for $\gamma^{AB} = \omega^A \omega^B$,

$$P = \zeta + \bar{\zeta} \tag{26}$$

with

$$\zeta = \frac{\mathrm{i}}{8\pi G} \oint \{(\Phi_{01} - \Psi_1)(\omega^0)^2 + (\Phi_{11} + \Lambda - \Psi_2)\omega^0\omega^1 + (\Phi_{21} - \Psi_3)(\omega^1)^2\}\boldsymbol{\mathcal{S}} \tag{27}$$

where $\boldsymbol{\mathcal{S}}$ is the surface-area 2-form on $\mathcal{S}$ and where the Ψ's. Φ's and Λ are (complex) Weyl tensor components, trace-free Ricci tensor components, and $R/24$, respectively. Since ζ is a quadratic form on $\mathbb{T}(\mathcal{S})$, the expression (27) defines an element

$$\boldsymbol{A} \in \mathbb{T}^*(\mathcal{S}) \circledS \mathbb{T}^*(\mathcal{S}), \tag{28}$$

which is the total *kinematic twistor* for the sources surrounded by $\mathcal{S}$.

To define the total rest-mass m surrounded by $\mathcal{S}$ we need a Hermitian squared norm (of signature $++--$) on the space $\mathbb{T}(\mathcal{S})$; then we find (by analogy with the flat-space twistor expressions for $\mathbb{M}$) that

$$m^2 = -\tfrac{1}{2}\|\boldsymbol{A}\|, \tag{29}$$

where the squared norm appearing in (29) is that naturally induced on the space (28). (An alternative suggestion made by Tod is to define m^4 in terms of $\det(\boldsymbol{A})$, which requires a volume element on $\mathbb{T}(\mathcal{S})$ rather than a norm.) A suggestion for the norm on $\mathbb{T}(\mathcal{S})$ is the average of

$$2\,\mathrm{Re}(\omega^0\bar{\pi}_0 + \omega^1\bar{\pi}_1) \tag{30}$$

over $\mathcal{S}$ ((30) being the standard constant norm of flat-space twistor theory), where $\bar{\pi}_0$ and $\bar{\pi}_1$ are the complex conjugates of

$$i\,\eth'\omega^1 - i\,\rho\omega^0 \quad \text{and} \quad i\,\eth\omega^0 - i\,\rho'\omega^1, \tag{31}$$

respectively. Here ρ and ρ' are real, describing the convergence of the null geodesics orthogonal to $\mathcal{S}$, i.e. the trace parts of the extrinsic curvature of $\mathcal{S}$ in $\mathcal{M}$. When $\mathcal{M}$ is conformally flat, and also in many other circumstances studied by Tod, (30) is constant over $\mathcal{S}$, so averaging is not necessary, and the resulting norm is conformally invariant. But in general (30) is not constant (and its average is not conformally invariant), so there remains some question as to the best definition of the norm.

Many other questions also remain. One of these is the non-negativeness of the right-hand side of (29), which would be expected if the 4-momentum part of the kinematics is to be time-like (as it should be, to make sense

physically). It would be desirable also to have a positivity proof for m at the quasi-local level, generalizing the Schoen–Yau–Witten arguments [19,24] for positivity of the ADM mass, etc. This would essentially follow if the right-hand side of (29) could be shown to be positive for sources satisfying the appropriate (dominant) energy condition.

REFERENCES

[1] R. Arnowitt, S. Deser and C. W. Misner, The dynamics of general relativity. In *Gravitation: An Introduction to Current Research*, edited by L. Witten, John Wiley and Sons, New York (1962).

[2] H. Bondi, Gravitational waves in general relativity. *Nature (Lond.)* **186** (1960), 535.

[3] H. Bondi, M. G. J. van der Burg and A. W. K. Metzner, Gravitational waves in general relativity. VII. Waves from axisymmetric isolated systems. *Proc. Roy. Soc. Lond. A* **269** (1962), 21–52.

[4] A. Einstein, *Sber. Preuss. Akad. Wiss.* (1918), 448.

[5] R. Geroch, A. Held and R. Penrose, A space-time calculus based on pairs of null directions. *J. Math. Phys.* **14** (1973), 874–881.

[6] R. Geroch and J. Winicour, Linkages in general relativity. *J. Math. Phys.* **22** (1981), 803–812.

[7] S. W. Hawking, Gravitational radiation in an expanding universe. *J. Math. Phys.* **9** (1968), 598–604.

[8] G. T. Horowitz and B. G. Schmidt, Note on gravitational energy. *Proc. Roy. Soc. Lond. A* **381** (1982), 215–224.

[9] A. Komar, Covariant conservation laws in general relativity. *Phys. Rev.* **113** (1959), 934–936.

[10] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon Press, Oxford (1962).

[11] R. Penrose, Asymptotic properties of fields and space-times. *Phys. Rev. Lett.* **10** (1963), 66–68.

[12] R. Penrose, General-relativistic energy flux and elementary optics. In *Perspectives in Geometry and Relativity*, edited by B. Hoffman, Indiana University Press, Bloomington, Ind. (1966).

[13] R. Penrose, Conserved quantities and conformal structure in general relativity. In *Relativity Theory and Astrophysics*, Vol. 1, *Relativity and Cosmology*, edited by J. Ehlers, American Mathematical Society, Providence, R.I. (1967).

[14] R. Penrose, Structure of space-time. In *Batelle Recontres, 1967 Lectures in Mathematics and Physics*, edited by C. M. deWitt and J. A. Wheeler, Benjamin, New York (1968).

[15] R. Penrose, Quasi-local mass and angular momentum in general relativity. *Proc. Roy. Soc. Lond. A* **381** (1982), 53–63.

[16] R. Penrose, Integrals for general-relativistic sources: a development from Maxwell's electromagnetic theory. In *AMS Maxwell Symposium Proceedings*, edited by M. Berger, American Mathematical Society, Providence, R.I. (1983).

[17] R. Penrose and W. Rindler, *Spinors and Space-Time*, Vols 1 and 2, Cambridge University Press, Cambridge (1983).

[18] R. K. Sachs, Gravitational waves in general relativity. VIII. Waves in asymptotically flat space-time. *Proc. Roy. Soc. Lond. A* **270** (1962), 103–126.

[19] R. Schoen and S.-T. Yau, Proof of the positive mass theorem II. *Commun. Math. Phys.* **79** (1981), 231–260.

[20] L. Tamburino and J. Winicour, Gravitational fields in finite and conformal Bondi frames. *Phys. Rev.* **150** (1966), 1039–1053.

[21] J. F. Taylor and P. M. McCullouch, Evidence for the existence of gravitational radiation from measurements of the binary pulsar PSR 1913 + 16. In *Ninth Texas Symposium on Relativistic Astrophysics*, edited by J. Ehlers, M. J. Perry and M. Walker, New York Academy of Sciences, New York (1980).

[22] K. P. Tod, Some examples of Penrose's quasi-local mass construction. *Proc. Roy. Soc. Lond. A* (to appear).

[23] J. Winicour, Angular momentum in general relativity. In *General Relativity and Gravitation*, Vol. 2, edited by A. Held, Plenum Press, New York (1980).

[24] E. Witten, A new proof of the positive energy theorem. *Commun. Math. Phys.* **80** (1981), 381–402.

150
'NEW IMPROVED' QUASI-LOCAL MASS, AND THE SCHWARZSCHILD SOLUTION

Twistor Newsletter **18** (1984). Also published in *Further Advances in Twistor Theory,* Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

 'New improved' quasi-local mass, and the Schwarzschild solution. *Twistor Newsletter* **18** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.15 'New improved' quasi-local mass, and the Schwarzschild solution *by R. Penrose* (**TN** 18, July 1984)

As has been pointed out by K.P.Tod in the previous article (§II.4.14), my original definition of quasi-local mass $m(\mathcal{S})$, for spacelike $\mathcal{S} \simeq S^2$ in a general space-time $\mathcal{M}$ (§II.4.2 and Penrose 1982) seems to run into difficulties in cases when the spacelike 2-surface $\mathcal{S}$ is *contorted* (i.e. not embeddable in a conformally flat space-time without changing its intrinsic or extrinsic curvatures). In view of the excellent results that K.P.Tod has earlier obtained with that definition (Tod 1983 and Shaw 1984) it would seem most unreasonable simply to abandon the quasi-local mass, but rather one should be seeking some simple modification of the original definition which gives agreement with these earlier results and which avoids the anomalies that can arise in contorted cases.

One of the more striking results that K.P.Tod had earlier obtained was that for any $\mathcal{S}$ drawn in any 3-surface of revolution $t = f(r)$ in the Schwarzschild solution $\mathcal{M}_m$, for mass m, the (original) quasi-local mass is always zero if $\mathcal{S}$ does not surround the source and is always m if $\mathcal{S}$ does surround it (just once). These are all *uncontorted* surfaces $\mathcal{S}$. The natural *conjecture* had been that the same result ought to hold if $\mathcal{S}$ does not lie in any $t = f(r)$ (and is therefore *contorted*). If one accepts the results for uncontorted $\mathcal{S}$'s in $\mathcal{M}_m$ then any other result in the contorted cases seems physically unreasonable. The result $(m(\mathcal{S}))^2 < 0$ for (certain) small contorted spheres, that K.P.Tod actually obtained (and was confirmed by R.Kelly) simply adds to this physical unreasonableness.

Just before the results of K.P.Tod's calculation were at hand, I had embarked on a seemingly promising-looking line of reasoning aimed at proving the above *conjecture*. But then K.P.Tod's results had appeared to invalidate this (then partly conjectural) line of argument. On re-examining my argument somewhat later I was surprised to find that it actually leads to a *modification* of my original definition for $A_{\alpha\beta}(\mathcal{S})$. R.Kelly and K.P.Tod then found that with such a modified definition, $m(\mathcal{S})$ indeed vanishes to 5th order for small contorted spheres in (e.g.) $\mathcal{M}_m$, as it presumably should.

My argument depended upon the existence of a *Killing spinor* (The sign 'Ξ' means that the LHS and the RHS are related by a constant factor which I haven't been bothered to sort out.)

$$X_{AB} \Xi \psi^{-1/3} \alpha_{(A} \beta_{B)},$$

where

$$\Psi_{ABCD} = \psi\alpha_{(A}\alpha_B\beta_C\beta_{D)}, \qquad \alpha_A\beta^A = 1$$

in $\mathcal{M}_m$ (cf. Walker & Penrose 1970) which satisfies

$$\nabla_{A'(A}X_{BC)} = 0.$$

This much holds also in any {22} vacuum (e.g. Kerr, NUT, C-metric). But I also require:

Property K. The field X^{AB}, restricted to $\mathcal{S}$, belongs to the symmetric tensor product of the 2-surface twistor space $\mathbb{T}(\mathcal{S})$ with itself.

The argument consists of two parts:

(1) Show that *if* property K holds, then the modified (i.e. 'new improved') quasi-local mass expression vanishes if $\mathcal{S}$ does not surround the source and gives the Schwarzschild mass m if $\mathcal{S}$ does surround it (just once). [The latter still presents problems.]

(2) Show that property K holds for the Schwarzschild solution.

Part (2) originated as a tentative conjecture. It is now known that his conjecture is false: property K must indeed fail for a general (contorted) $\mathcal{S}$ in the Schwarzschild solution (as follows from a calculation by N.M.J.Woodhouse, Woodhouse (1987), cf. §II.4.14 and Kelly, Tod & Woodhouse (1986)Nevertheless, I beleive that there is some merit in giving my argument for (1) here and for outlining my attempted argument aimed in the direction of (2). Perhaps there are pointers here towards finding out what is *really* going on.

Part (1). We work in $\mathbb{T}^*(\mathcal{S})$, which is a flat (dual) twistor space, and consider the twistor contour integral

$$A_{\alpha\beta} = \oint \frac{W_\alpha W_\beta \, d^4W}{(K^{\rho\sigma}W_\rho W_\sigma)^3}. \qquad (A)$$

Here $K^{\alpha\beta}$ is a twistor in $\mathbb{T}(\mathcal{S}) \odot \mathbb{T}(\mathcal{S})$ whose primary part agrees with X^{AB} on $\mathcal{S}$. For this we need property K. Also we need $\mathcal{S}$ to be embedded in $\mathbb{M}(\mathcal{S})$—the complex compact Minkowski space associated with $\mathbb{T}^*(\mathcal{S})$ (i.e. with $\mathbb{T}(\mathcal{S})$). The embedding is easily achieved, where we identify each point P of $\mathcal{S}$ with the linear 2-space in $\mathbb{T}^*(\mathcal{S})$ whose elements take the local form $(\lambda_A, 0)$ at P (and so annihilate the 2-surface twistor fields ω^A which vanish at P). We also need a choice of

twistor $\varepsilon^{\alpha\beta\gamma\delta}$ to define the d^4W:

$$d^4W = \tfrac{1}{24}\, dW_\alpha \wedge dW_\beta \wedge dW_\gamma \wedge dW_\delta \varepsilon^{\alpha\beta\gamma\delta}. \tag{B}$$

To perform the integral (A) we can introduce a basis $\delta_0^\alpha, \delta_1^\alpha, \delta_2^\alpha, \delta_3^\alpha$ for $\mathbb{T}(\mathcal{I})$ and take components with respect to it:

$$\lambda_0 = W_0 = W_\alpha \delta_0^\alpha, \quad \lambda_1 = W_1 = W_\alpha \delta_1^\alpha, \quad \mu^{0'} = W_2 = W_\alpha \delta_2^\alpha, \quad \mu^{1'} = W_3 = W_\alpha \delta_3^\alpha. \tag{C}$$

Note that each of $\delta_0^\alpha, \cdots, \delta_3^\alpha$ is a 2-surface twistor and so has a description in terms of '$(\omega^0, \omega^1, \pi_{0'}, \pi_{1'})$' which varies over $\mathcal{I}$ according to the standard 2-surface twistor equations on $\mathcal{I}$. If at one point Q on $\mathcal{I}$ we choose

$$\varepsilon^{\alpha\beta\rho\sigma} = 24\, \delta_0^{[\alpha}\delta_1^\beta\delta_2^\rho\delta_3^{\sigma]}, \qquad \text{so } \varepsilon^{0123} = 1 \tag{D}$$

then this equation will hold at all points of $\mathcal{I}$. But in terms of the standard local '$(\omega^0, \cdots, \pi_{1'})$' on $\mathcal{I}$ we shall find that the components of the spinor parts $(\cdots, \varepsilon^{AB}\varepsilon_{R'S'}, \cdots)$ may vary from point to point on $\mathcal{I}$. If $\mathcal{I}$ is uncontorted then these components are in fact constant over $\mathcal{I}$. But in the general case they vary. There is just one overall factor to be concerned with—a scalar field ν on $\mathcal{I}$. Then the components of these spinor parts are just $\pm\nu$ and 0 (This ν is just the $\tilde{\Omega}\Omega^{-1}$ of Tod's article). Tod conjectures that ν has constant modulus over $\mathcal{I}$. This would be important. Then fixing $\nu = 1$ at Q, say, we should have ν of *unit modulus* over $\mathcal{I}$.

Reverting, for the moment, to twistor components with respect to the 'constant' twistor frame of (C), we can write (with component indices underlined)

$$\mu^{\underline{A}'} = -ix^{\underline{A}\underline{A}'}\lambda_{\underline{A}} \tag{E}$$

where $x^{\underline{A}\underline{A}'}$ are Minkowski cooordinates for $\mathbb{M}(\mathcal{I})$. Then we can write

$$d^4W = d^2\lambda \wedge d^2\mu = -\lambda_{\underline{A}}\lambda_{\underline{B}}\mathrm{d}x^{\underline{A}0'} \wedge dx^{\underline{B}1'} \wedge d\lambda_0 \wedge d\lambda_1 \tag{F}$$

and perform the λ-integral first, in (A), to obtain

$$A_{\alpha\beta}Z^{\alpha}Z^{\beta} = \oint \frac{\omega^{\underline{A}}\omega^{\underline{B}}\lambda_{\underline{A}}\lambda_{\underline{B}}\lambda_{\underline{C}}\lambda_{\underline{D}}}{(X^{\underline{P}\underline{Q}'}\lambda_{\underline{P}}\lambda_{\underline{Q}})^3}\, dx^{\underline{C}0'} \wedge dx^{\underline{D}1'} \wedge d\lambda_0 \wedge d\lambda_1$$

$$= \oint \Psi_{\underline{ABCD}}\omega^{\underline{A}}\omega^{\underline{B}}\ dx^{\underline{C}0'} \wedge dx^{\underline{D}1'} \qquad \text{(G)}$$

since

$$\Psi_{ABCD} = \oint \frac{\lambda_A\lambda_B\lambda_C\lambda_D}{(X^{PQ}\lambda_P\lambda_Q)^3}\, d\lambda_0 \wedge d\lambda_1 \qquad \text{[at each fixed point of } \mathcal{S}\text{]} \qquad \text{(H)}$$

(by direct calculation, if desired). Now the expression in the final right hand side of (G) is just the standard (original) quasi-local mass expression (in the case of vacuum) *except* for the fact that it is the differentials $dx^{\underline{C}0'} \wedge dx^{\underline{D}1'}$, for the embedding of $\mathcal{S}$ in $\mathbb{M}(\mathcal{S})$, which appear, rather than the local surface element for $\mathcal{S}$. In order to rewrite this integral in terms of this local surface element, we now return to descriptions in terms of the *local* (o^A, ι^A) spinor frame. This entails that the factor ν now appears in the integral (G).

Guessing that this applies generally for an arbitrary space-time which need not be vacuum, we are led to the '*new improved*' *quasi-local mass-angular momentum object*

$$A_{\alpha\beta}Z^{\alpha}Z^{\beta} = \frac{-i}{4\pi G}\oint \nu\Big\{(\Psi_1 - \Phi_{10})\omega^0\omega^0 + 2(\Psi_2 - \Phi_{11} - \Lambda)\omega^0\omega^1 + (\Psi_3 - \Phi_{21})\omega^1\omega^1\Big\}\,\underline{\mathcal{S}} \qquad \text{(I)}$$

where $\underline{\mathcal{S}}$ is the surface area element. If ν is indeed of unit modulus, in accordance with K.P.Tod's conjecture, then (I) seems well-enough defined for its necessary purposes. Otherwise some procedure would be needed to fix $|\nu|$, e.g. normalizing so that the average $|\nu|$ over $\mathcal{S}$ is unity.

Now, by the methods of standard flat-space twistor theory (cf. Penrose & MacCallum 1973) we can evaluate (A), and find

$$A_{\alpha\beta} = \frac{K^{(-1)}{}_{\alpha\beta}}{\sqrt{detK^{\rho\sigma}}} \qquad \text{or} \qquad 0 \qquad \text{(J)}$$

according as: $\mathcal{S}$ surrounds the source (once); or else can be shrunk to a point without crossing the source region. To compute the mass from (J) when the source *is* linked we still have some problems. Perhaps Tod's method of forming a determinant and then taking $|\cdots\cdots|^{1/4}$ may be the best. But this seems to need Tod's conjecture.

Part (2). (Outline of my original attempt to prove property K for Schwarzchild.) We need a usable method of distinguishing Schwarschild case from C-metric (since we *know* $A_{\alpha\beta} \neq 0$ in case

 'New improved' quasi-local mass, and the Schwarzschild solution. *Twistor Newsletter* **18** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

when $\mathcal{I}$ lies on $\mathcal{I}^+$ above the 'bullet holes' of the C-metric) and also from Kerr case (since K.P.Tod and W.T.Shaw have shown that property K is *false* for Kerr). Now $\mathcal{M}_m$ has 3-rotational Killing vectors x^a, y^a, z^a, with

$$X_{AB}\tilde{X}_{A'B'} = x_{AA'}x_{BB'} + y_{AA'}y_{BB'} + z_{AA'}z_{BB'} + \lambda\varepsilon_{AB}\varepsilon_{A'B'} \qquad \text{(K)}$$

(Here $\tilde{X}_{A'B'}$ is the complex conjugate of X_{AB}, but will need to be 'freed' from it shortly.) The tangential parts of

$$\nabla_{A'(A}X_{BC)} = 0, \quad \nabla_{A(A'}\tilde{X}_{B'C')} = 0, \quad \nabla_{(A|(A'}x_{B')|B)} = 0, \ \cdots y, \ \cdots z \cdots$$

give

$$\eth X_{00} + 2\sigma X_{10} = 0, \ \eth' X_{11} + 2\sigma' X_{01} = 0, \ \cdots \tilde{X} \cdots, \ \cdots, \ \eth' x_{01'} + \sigma x_{11'} + \bar{\sigma}' x_{00'} = 0, \ \cdots, \ y, \cdots z \cdots,$$

which provide 10 differential equations relating derivatives of the 10 'nice' components

$$X_{00}, \ X_{11}, \ \tilde{X}_{0'0'}, \ X_{1'1'}, \ x_{01'}, \ x_{10'}, \ y_{01'}, \ y_{10'}, \ z_{01'}, \ z_{10'}$$

to the 9 'nasty' components

$$X_{01}, \ \tilde{X}_{0'1'}, \ x_{00'}, \ y_{11'}, \ y_{00'}, \ y_{11'}, \ z_{00'}, \ z_{11'}, \ \lambda.$$

From (K) we have 2 equations

$$X_{00}\tilde{X}_{1'1'} = x_{01'}^2 + y_{01'}^2 + z_{01'}^2,$$
$$X_{11}\tilde{X}_{0'0'} = x_{10'}^2 + y_{10'}^2 + z_{10'}^2$$

relating only nice components and 6 equations

$$X_{00}\tilde{X}_{0'1'} = x_{01'}x_{00'} + y_{01'}y_{00'} + z_{01'}z_{10'} \quad \text{etc.}$$

which relate nasty ones to nice ones. Two more equations on nasty components are provided by

$$X_{01}{}^2 - X_{00}X_{11} = \lambda = \tilde{X}_{0'1'} - \tilde{X}_{0'0'}\tilde{X}_{1'1'} \qquad \text{(L)}$$

and we also have, from the skew parts of (K),

$$X_{01}\bar{X}_{0'1'} = x_{01'}x_{10'} + y_{01'}y_{10'} + z_{01'}z_{10'} - \lambda = x_{00'}x_{11'} + y_{00'}y_{11'} + z_{00'}z_{11'} + \lambda \qquad (\mathrm{M})$$

though only *one* of these is independent of (L), in effect. In principle, we could use these $6+2+1$ equations to express all the nasty components in terms of nice ones and substitute back into the 10 differential equations. The question is: how big is the solution space? The idea was to try to show that it is no bigger than in the flat case, where we know that *all* such X's, $\bar{X}$'s, x's, y's, z's do belong to their proper $\mathbf{T}\odot\mathbf{T}$, $\mathbf{T}^*\odot\mathbf{T}^*$, $\mathbf{T}\otimes\mathbf{T}^*$, $\mathbf{T}\otimes\mathbf{T}^*$, $\mathbf{T}\otimes\mathbf{T}^*$. Then since, in $\mathcal{M}_m$, all of X_{AB}, $\bar{X}_{A'B'}$, $x_{AA'}$, $y_{AA'}$, $z_{AA'}$ actually *do* satisfy all the equations, it must follow that they belong to the required spaces $\mathbf{T}(\mathcal{I})\odot\mathbf{T}(\mathcal{I})$, etc. To see how big the solution space actually is, the idea was to consider the 'canonical' case first, when $\mathcal{I}$ is spherically symmetrically situated, so $\sigma = 0$, $\sigma' = 0$. The differential equations decouple and can be solved explicitly for the nice components. From (L) and (M) we can find X_{01} and $\bar{X}_{0'1'}$ in terms of nice components and then solve for the remaining nasty ones. The hope had been that a general argument could be invoked to show that no new solutions appear on perturbation. Howver we know now that this does not work.

Thanks to K.P.Tod, R.M.Kelly, W.T.Shaw and N.M.J.Woodhouse.

References

Penrose, R. (1982) Proc. Roy. Soc. Lond. **A381**, 53-63.

Penrose, R. & MacCallum, M.A.H. (1973) Phys. Repts. **6C**, no. 4, 241.

Shaw, W.T. (1984) Proc. Roy. Soc. Lond., **A390**, 191-215.

Tod, K.P. (1983) Some examples of Penrose's quasi-local mass construction, Proc. Roy. Soc. Lond. **A388** 457.

Kelly, R., Tod, K.P. and Woodhouse, N.M.J. (1986) Quasi-local mass for small surfaces, Class. & Quant. Grav. **3**, 1151-67).

Walker, M. & Penrose, R. (1970) Comm. Math. Phys., **18**, 265-274

Woodhouse, N.M.J. (1987) Ambiguities in the definition of quasi-local mass, Class. Quant. Grav., **4**, L121-123

151
ON BRYANT'S CONDITION FOR HOLOMORPHIC CURVES IN CR SPACE

Twistor Newsletter **18** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 On Bryant's condition for holomorphic curves in CR space. *Twistor Newsletter* **18** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.3.9 **On Bryant's Condition for Holomorphic Curves in CR-Spaces** *by R. Penrose* (TN 18, June 1984)

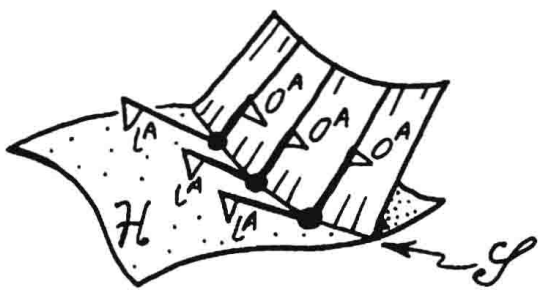
I am concerned here with 5-dimensional CR-manifolds and the conditions on such a manifold $\mathcal{C}$ that it be the space of projective null twistors for a hypersurface twistor space: $\mathcal{C} = \mathbb{P}\mathcal{N}(\mathcal{H})$.

$\mathcal{H}$ is a spacelike hypersurface in a (real) space-time manifold $\mathcal{M}$ (where neither need be assumed analytic). Each point of $\mathcal{H}$ gives rise to a holomorphic curve ($\cong S^2$) in $\mathcal{C}$, so $\mathcal{C}$ is fibred by this 3-real-parameter family of holomorphic curves. When $\mathcal{M}$ is conformally flat, this extends to a 4-real-parameter family of holomorphic curves (corresponding to the points of $\mathcal{M}$). LeBrun has investigated cases of *non*-conformally flat $\mathcal{M}$'s where $\mathbb{P}\mathcal{N}(\mathcal{H})$ contains holomorphic curves other than those corresponding to points of $\mathcal{H}$; this does *not* happen for the generic $\mathbb{P}\mathcal{N}(\mathcal{H})$.

Robert Bryant has shown that if $\mathcal{C}$ is merely given as a $(+-)$-Levi-signature CR-5-manifold, then a necessary and sufficient condition that it be (locally) foliated by holomorphic curves is that a certain quartic polynomial P, constructed from the Chern-Moser 'S-tensor' should have a *repeated root.* Thus it follows that every $\mathbb{P}\mathcal{N}(\mathcal{H})$ must satisfy Bryant's repeated root condition. A rough dimension count indicated that this condition would not be sufficient to characterize $\mathbb{P}\mathcal{N}(\mathcal{H})$'s and a complete characterization of such spaces was left open.

I shall give an argument to show that in any $\mathbb{P}\mathcal{N}(\mathcal{H})$, P must actually have (at least) a *triple* root.

Consider a point X of $\mathcal{C} = \mathbb{P}\mathcal{N}(\mathcal{H})$. The roots of P at X correspond to the possible choices of 2-surface element E at X which, 'to third order', could be tangent to a holomorphic curve, i.e., E is tangent to a smooth 2-surface $\mathcal{E}$, in $\mathcal{C}$, where $\mathcal{E}$ is holomorphic to 3rd order at X . Corresponding to $\mathcal{E}$, there is a smooth 2-surface $\mathscr{S}$ in $\mathcal{H}$ and a field of null directions (in $\mathcal{M}$) over $\mathscr{S}$—or else $\mathscr{S}$ degenerates at X (as would be the case when $\mathcal{E}$ actually corresponds to a point of $\mathcal{H}$). When $\mathscr{S}$ is indeed a smooth 2-surface (at X), the condition for holomorphicity at X is that: (a) the specified null directions at $\mathscr{S}$ are normals/tangents to a null hypersurface $\mathcal{Q}$ in $\mathcal{M}$ which intersects $\mathcal{H}$ locally in $\mathscr{S}$ (condition of 'abreastness' of the null rays, arising since the 1-tangents to $\mathcal{E}$ lie in the holomorphic tangent space to $\mathcal{C}$ at X); and (b) that the complex shear σ of $\mathcal{Q}$ at X, together with its 1st and 2nd derivatives in directions tangent to $\mathscr{S}$, should vanish.



I use the compacted spin-coefficient (GHP) formalism. The flagpole of o^A at each point of $\mathscr{S}$ is chosen in the specified null direction; the flagpole of ι^A is taken in the other null direction orthogonal to $\mathscr{S}$. We have

$$\eth\rho - \eth'\sigma = (\rho - \bar{\rho})\tau + (\bar{\rho}' - \rho')\kappa - \Psi_1 + \Phi_{01}.$$

Apply $\eth'$, then take the difference between the resulting expression and its complex conjugate, using commutator relations for $\eth$, $\eth'$ to obtain:

$$\eth'^2\sigma - \eth^2\bar{\sigma} = \eth'\Psi_1 - \eth'\Phi_{01} - \eth\Psi_1 + \eth\Phi_{10} + \rho(\sigma\sigma' - \bar{\sigma}\bar{\sigma}' - \Psi_2 + \bar{\Psi}_2).$$

For a 'holomorphic curve to third order' we require $\sigma = 0, \ldots, \eth^2\sigma = 0$, whence

$$\rho = \frac{\mathrm{Im}(\eth'\Psi_1 - \eth'\Phi_{01})}{\mathrm{Im}\,\Psi_2}.$$

 On Bryant's condition for holomorphic curves in CR space. *Twistor Newsletter* **18** (1984). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

Note that in the *general case*, when $\mathrm{Im}\,\Psi_2 \neq 0$, there is only one such ρ. Now ρ $(= \bar{\rho})$ and σ $(= 0)$ fix the direction of E at X, so in the general case we get just *one* root of P corresponding to smooth $\mathscr{S}$. It is easy enough to see that (because $\sigma = 0$ at X) the 'degenerate' cases of non-smooth $\mathscr{S}$ at X arise only when the holomorphic curve is one arising from a point of $\mathcal{H}$. These correspond to Bryant's repeated root of P. Since P is quartic, the only possible coincidence schemes are $\{211\}$, $\{22\}$, $\{31\}$, $\{4\}$, $\{-\}$. Of these, only $\{22\}$ and $\{31\}$ allow just a single root for smooth $\mathscr{S}$. But, as LeBrun points out, $\{2,2\}$ is ruled out in the *general* case since we would then have a second family of holomorphic curves. Hence, $\{31\}$ *is*, indeed, the general case (with $\{4\}$ or $\{-\}$ in special situations). Thanks to Peter Thompson and Claude LeBrun.

References

Bryant, R.L. (1982) Trans. Am. Maths. Soc., **272**, No 1, July, p. 203.

Geroch, R., Penrose, R. & Held, A. (1973) J. Math. Phys. **14**, 874-881.

Penrose, R. & Rindler, W. (1984) Spinors and Space-time, Vol. 1, Cambridge University Press, Cambridge.

152
INTEGRALS FOR GENERAL-RELATIVISTIC SOURCES: A DEVELOPMENT FROM MAXWELL'S ELECTROMAGNETIC THEORY

The sesquicentennial symposium (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. North-Holland, Amsterdam, 1984

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

INTEGRALS FOR GENERAL-RELATIVISTIC SOURCES: A DEVELOPMENT FROM MAXWELL's ELECTROMAGNETIC THEORY.

Roger PENROSE
Mathematical Institute, Oxford, U.K.

Maxwell's pioneering work on electromagnetism produced a paradigm—a field theory of astonishing accuracy and of remarkable mathematical elegance and simplicity. One direct consequence of this theory was that the charge surrounded by a closed 2-surface S can be computed by integrating field components over S. In general relativity, by introducing a twistor concept associated with S (taken to be spacelike topological 2-sphere), the gravitational field near S (spin 2) may be reduced to the Maxwell chase (spin 1). The resulting 'charge' integrals provide suggestive definitions for the total mass-momentum and angular momentum surrounded by S.

1. Maxwell's equations

I feel priviledged to have this opportunity to pay homage to James Clerk Maxwell, who must surely rate as one of the very greatest mathematical physicists of all time. And among his fundamental contributions to the development of physics, his electromagnetic theory shines out as the one of greatest significance. I think that it is hard for us to appreciate, for our vantage point of the late 20th century, when the concept of field theory has gained such broad and seemingly 'obvious' acceptance, how revolutionary Maxwell's ideas actually were. With the notable exception of Faraday before him, no other major physicist of his day had apparently regarded the concept of 'field' as anything more than a convenient mathematical auxilliary to the prevailing point-particle-action-at-a-distance view of physical reality. The idea that 'disembodied fields' can propagate through empty space carrying energy as they go, was as

startling and revolutionary an idea at the time as radio is commonplace to us today. (For an elegant readable account of Maxwell's life and achievements, see Tolstoy, 1981.)

Let us examine how the description of Maxwell's equations has evolved over the years. First, there is the form in which Maxwell actually first gave them [13] (where I set the permeability $\mu = 1$ and write the velocity of light c for Maxwell's E, the more usual $c^{-2}\rho$ for his charge density e, and the now standard ∂ for his partial derivative d):

$$\begin{aligned} &\frac{\partial\alpha}{\partial x}+\frac{\partial\beta}{\partial y}+\frac{\partial\gamma}{\partial z}=0, && \frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}=4\pi\rho \\ &\frac{\partial Q}{\partial z}-\frac{\partial R}{\partial y}=\frac{\partial\alpha}{\partial t}, && \frac{\partial\gamma}{\partial y}-\frac{\partial\beta}{\partial z}=\frac{1}{c^2}\frac{\partial P}{\partial t}+4\pi p \\ &\frac{\partial R}{\partial x}-\frac{\partial P}{\partial z}=\frac{\partial\beta}{\partial t}, && \frac{\partial\alpha}{\partial z}-\frac{\partial\gamma}{\partial x}=\frac{1}{c^2}\frac{\partial Q}{\partial t}+4\pi q \\ &\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}=\frac{\partial\gamma}{\partial t}, && \frac{\partial\beta}{\partial x}-\frac{\partial\alpha}{\partial y}=\frac{1}{c^2}\frac{\partial R}{\partial t}+4\pi r \end{aligned} \tag{1.1}$$

together with the 'equation of continuity' (charge conservation)

$$\frac{\partial p}{\partial x}+\frac{\partial q}{\partial y}+\frac{\partial r}{\partial z}+\frac{1}{c^2}\frac{\partial\rho}{\partial t}=0, \tag{1.2}$$

where (α, β, γ) are the components of the magnetic field vector $\boldsymbol{B}$, where (P, Q, R) are those of the electric field vector $\boldsymbol{E}$, and (p, q, r) those of the electric current vector j. The vector notation of Gibbs and Heaviside enabled the appearance of these equations to be simplified to

$$\begin{aligned} &\operatorname{div}\boldsymbol{B}=0 && \operatorname{div}\boldsymbol{E}=4\pi\rho \\ &-\operatorname{curl}\boldsymbol{E}=\frac{\partial\boldsymbol{B}}{\partial t} && \operatorname{curl}\beta=\frac{1}{c^2}\frac{\partial\boldsymbol{E}}{\partial t}+4\pi j \end{aligned} \tag{1.3}$$

together with

$$\operatorname{div} j+\frac{1}{c^2}\frac{\partial\rho}{\partial t}=0. \tag{1.4}$$

Minkowski's (1908) 4-dimensional formulation of special relativity

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

subsequently led to a further simplification of the equations to[1]

$$\nabla_{[a}F_{bc]} = 0, \qquad \nabla_a F^{ab} = 4\pi J^b \tag{1.5}$$

with

$$\nabla_a J^a = 0\,, \tag{1.6}$$

where

$$F_{ab} = \begin{bmatrix} 0 & P & Q & R \\ -P & 0 & -c\gamma & c\beta \\ -Q & c\gamma & 0 & -c\alpha \\ -R & -c\beta & c\alpha & 0 \end{bmatrix} \qquad J^a = \begin{bmatrix} \rho \\ cp \\ cq \\ cr \end{bmatrix} \tag{1.7}$$

and

$$\nabla_a = \frac{\partial}{\partial x^a} \quad \text{with } x^0 = ct,\ x^1 = x,\ x^2 = y,\ x^3 = z\,, \tag{1.8}$$

indices being raised and lowered in the standard way using the Minkowski (pseudo-)metric tensor components

$$g_{ab} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = g^{ab}\,. \tag{1.9}$$

Nowadays, mathematicians usually prefer to use the index-free differential form notation of É. Cartan in which Maxwell's equations take the remarkably simple form

$$\mathrm{d}\boldsymbol{F} = 0, \qquad \mathrm{d}\boldsymbol{F}^* = 8\pi \boldsymbol{J}^*\,, \tag{1.10}$$

with

$$\mathrm{d}\boldsymbol{J}^* = 0\,. \tag{1.11}$$

Here the Maxwell field 2-form is

$$\boldsymbol{F} = F_{ab}\,\mathrm{d}x^a \wedge \mathrm{d}x^b \tag{1.12}$$

[1]Summation over repeated indices is assumed throughout. Square brackets around indices denote anti-symmetrization; likewise, round brackets around indices denote symmetrization.

with Hodge dual

$$\boldsymbol{F}^* = \tfrac{1}{2} e_{\mathbf{abcd}} F^{\mathbf{cd}}\, \mathrm{d}x^{\mathbf{a}} \wedge \mathrm{d}x^{\mathbf{b}}\,, \tag{1.13}$$

where the alternating tensor symbol $e_{\mathbf{abcd}}$ satisfies $e_{\mathbf{abcd}} = e_{[\mathbf{abcd}]}$ and $e_{0123} = 1$, and the charge-current 1-form is

$$\boldsymbol{J} = J_{\mathbf{a}}\, \mathrm{d}x^{\mathbf{a}}$$

with Hodge dual

$$\boldsymbol{J}^* = \tfrac{1}{6} e_{\mathbf{abcd}} \boldsymbol{J}^{\mathbf{d}}\, \mathrm{d}x^{\mathbf{a}} \wedge \mathrm{d}x^{\mathbf{b}} \wedge \mathrm{d}x^{\mathbf{c}}\,. \tag{1.14}$$

The simplicity of (1.10) (and of the consistency condition (1.11)) is very striking. This, to some degree, is to the credit of developments in modern mathematical notation and concepts. But the simplicity is also inherent in the Maxwell equations themselves—reflecting a harmony in nature itself. Indeed, the development of the very mathematical ideas involved in the calculus of differential forms historically owes much to the study of the elegant properties of electromagnetic phenomena.

A central theorem of that calculus is the result that may be referred to as the *fundamental theorem of exterior calculus* (the famous formula due to Ostrogradski, Gauss, Green, Kelvin, Stokes, Cartan etc.) which, in its general form, asserts

$$\int_{\Sigma} \mathrm{d}\boldsymbol{\alpha} = \oint_{\partial\Sigma} \boldsymbol{\alpha}\,. \tag{1.15}$$

where $\boldsymbol{\alpha}$ is some p-form and $\partial\Sigma$ the (smooth oriented) boundary of a compact oriented $(p+1)$-dimensional region Σ in some n-manifold ($n > p > 0$). This theorem can be applied at various levels in Maxwell theory. Most obviously, when $p = 3$, $n = 4$, the theorem directly expresses conservation of electric charge (applied to $\mathrm{d}\boldsymbol{J}^* = 0$):

$$\oint_{\partial\Sigma} \boldsymbol{J}^* = 0\,. \tag{1.16}$$

Here (see Fig. 1) we may envisage $\partial\Sigma$ to be made up of three pieces

$$\partial\Sigma = A_2 - A_1 + H \tag{1.17}$$

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

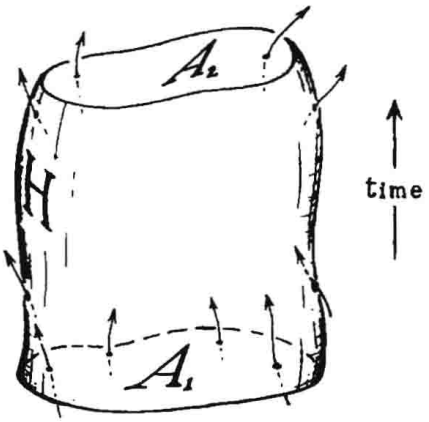


Fig. 1

(where the signs are to deal with the orientations consistently). A_1 and A_2 are two (positively future-oriented) spacelike 3-volumes in space-time, representing bounded spatial volumes 'at different times', so that $\int_{A_i} \boldsymbol{J}^*$ measures the total charge in A_i $(i = 1, 2)$, and H is a timelike region (3-volume) representing the 'history' of the *boundary* of the spatial region as it evolves from A_1 to A_2, so that $\int_H \boldsymbol{J}^*$ measures the net total outward flux of charge across this boundary. The relation

$$\int_{A_2} \boldsymbol{J}^* = \int_{A_1} \boldsymbol{J}^* - \int_H \boldsymbol{J}^* \tag{1.18}$$

obtained by substituting (1.17) into (1.16), expresses the fact that the charge lost from a spatially compact region is precisely that which flows across its boundary.

The formula (1.15) also directly applies to Maxwell theory when $p = 2$, $n = 4$ where now we choose either $\boldsymbol{\alpha} = \boldsymbol{F}^*$ or $\boldsymbol{\alpha} = \boldsymbol{F}$. Let us take $\partial\Sigma$ to be a spacelike 2-surface S surrounding (i.e. linking) some world-tube of charge (Fig. 2). Then the integral

$$e = \frac{1}{8\pi} \oint \boldsymbol{F}^* \tag{1.19}$$

measures the total charge within S, since if Σ is any compact 3-volume spanning S (so $S = \partial\Sigma$), the total charge on Σ, which is $\int_\Sigma \boldsymbol{J}^*$, is equal to (1.19) by (1.15) and (1.10). Note that it does not matter *which* 3-volume

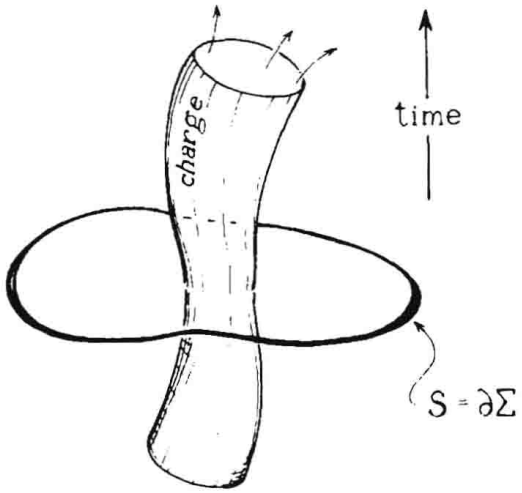


Fig. 2.

Σ we choose to span S, this fact being a special case of the charge conservation (1.18) (when $H = 0$). Thus we can think of the charge 'surrounded by' S as a property simply of S and the electromagnetic field *at* S—we do not need to examine the field or charge-current elsewhere than at S.

Similarly we can perform the corresponding integral to (1.19) but with $\boldsymbol{F}$ in place of $\boldsymbol{F}^*$. By (1.10) and (1.15) we find

$$\frac{1}{8\pi}\oint_S \boldsymbol{F} = 0\,, \tag{1.20}$$

which expresses the fact that the total *magnetic* charge is always zero in Maxwell theory (absence of magnetic monopoles). We can also use (1.15) at the $p = 1$ level when a 4-potential 1-form $\boldsymbol{\Phi}$ is introduced, with $\boldsymbol{F} = \mathrm{d}\boldsymbol{\Phi}$, but I shall not pursue this.

My main purpose, here, will be to exhibit a corresponding formula to (1.19) which applies in Einstein's general relativity and expresses the mass (-momentum and angular momentum) similarly surrounded (i.e. linked) by a 2-surface. Einstein's theory is the field theory that does for gravity that Maxwell's theory does for electromagnetism. But, for various reasons, Einstein's theory is considerably more difficult to handle than Maxwell's. In particular, whereas the *field* in Maxwell's theory does not act as its own source (i.e. the electromagnetic field is itself uncharged), this is not so in Einstein's theory. The gravitational field does

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

(in a strange non-local way) possess energy and therefore mass, while mass itself acts as source for the gravitational field. Thus Einstein's theory turns out to be a highly *non-linear* theory, in contradistinction to the linearity of Maxwell's theory.

Before exhibiting my promised formula it will be necessary to digress somewhat, in two different but related directions. I first consider how both Maxwell's theory and the weak-field (linear) limit of Einstein's (when source-free) are both part of a general pattern, being effectively the spin 1 and spin 2 cases of the general spin[2] massless free-field equations. Then I introduce the concept of a *twistor*, which supplies the link (in steps of $\frac{1}{2}$) between massless fields of different spins. When applied in a curved-space context this enables the gravitational (spin 2) version of (1.19) to be obtained from its electromagnetic (spin 1) counterpart, using twistors, in two steps, to push the spin down from two to one.

2. Massless fields of arbitrary spin

Since I shall need to discuss half-integral as well as integral spins, it will be necessary to introduce some spinor notation. For this, we express the Minkowski components V^0, V^1, V^2, V^3 of a *world vector* $\boldsymbol{V}$ (i.e. a possibly complexified tangent vector to space-time) in the matrix form

$$V^{AA'} = \begin{bmatrix} V^{00'} & V^{01'} \\ V^{10'} & V^{11'} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} V^0 + V^1 & V^2 + \mathrm{i}V^3 \\ V^2 - \mathrm{i}V^3 & V^0 - V^1 \end{bmatrix}, \tag{2.1}$$

which is Hermitian whenever $\boldsymbol{V}$ is real. When $\boldsymbol{V}$ is a (complex) *null* vector, i.e.

$$V^a V^b g_{ab} = 0\,, \tag{2.2}$$

[2]The concept of 'spin' has a clear-cut geometrical meaning for massless fields. Let us envisage a plane wave of the field and then examine the effect of rotating it through an angle θ about the propagation direction. If the spin is j ($= \frac{1}{2}, 1, 2, \ldots$) then the field will be restored to its original value if $\theta = 2\pi/j$; moreover this is the smallest (positive) angle with this property. (When $j = 0$, any θ has this property.) Note that for a Maxwell field we need $\theta = 2\pi$ whereas a gravitational wave, being a wave of tidal curvature, is restored when $\theta = \pi$. Fractional spin requires two-valued spinor fields, which are sent to their negatives if $\theta = 2\pi$. The relation between this j and the helicity s described in §2 is $j = |s|$. For a (positive frequency) field of helicity s, rotation through θ about the propagation direction multiplies the field by $e^{is\theta}$.

the matrix in (2.1) has rank 1 (or 0) and so has the form

$$V^{AA'} = \xi^A \eta^{A'} \tag{2.3}$$

We think of $\xi^A = (\xi^0, \xi^1)$ as being the (complex) components of a *spin-vector* $\boldsymbol{\xi}$ (or reduced spinor, in the sense of Cartan) and $\eta^{A'} = (\eta^{0'}, \eta^{1'})$ the components of a *conjugate spin-vector* $\boldsymbol{\eta}$. At each point of space-time the spin-vectors constitute a complex two-dimensional vector space referred to as spin-space $\mathbb{S}$ and the conjugate spin-vectors similarly constitute the *conjugate* spin-space $\bar{\mathbb{S}}$, at that space-time point. Under the operation of complex conjugation (which leaves real world vectors invariant), $\mathbb{S}$ and $\bar{\mathbb{S}}$ are interchanged. The essential content of (2.1) is that the real four-dimensional space $\mathbb{V}$ of world-vectors, at the space-time point in question, is the Hermitian part of the tensor product of $\mathbb{S}$ with $\bar{\mathbb{S}}$

$$\mathbb{V} = \mathrm{Herm}(\mathbb{S} \otimes \bar{\mathbb{S}}), \tag{2.4}$$

where we *identify* complex world vectors with their spinor counterparts according to

$$\mathbb{V} \oplus i\mathbb{V} = \mathbb{S} \otimes \bar{\mathbb{S}}. \tag{2.5}$$

The action of $\mathrm{SL}(2, \mathbb{C})$ on $\mathbb{S}$, accompanied by the complex conjugate action on $\bar{\mathbb{S}}$, provides the action of the restricted Lorentz group on $\mathbb{V}$:

$$\mathrm{SL}(2, \mathbb{C}) \xrightarrow{2\text{–}1} \mathrm{O}\uparrow(1, 3). \tag{2.6}$$

It is virtually essential, when performing detailed manipulations with spinors, that the appropriate index notation be at hand. This does not require that a basis be chosen, since the indices can simply be *abstract markers* and need not take numerical values in the conventional way. Whereas I have been using bold upright letters to denote conventional numerical component indices, I shall henceforth use standard lightface italic indices to denote such abstract markers. One advantage of this abstract index notation is that we can simply *identify the markers*

$$a = AA', \qquad b = BB', \ldots, \qquad z = ZZ', \qquad a_0 = A_0A_0', \ldots \tag{2.7}$$

and write

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

$$V^a = V^{AA'}, \qquad V^b = V^{BB'}, \ldots \tag{2.8}$$

for different abstract-indexed versions of the same world vector $\boldsymbol{V}$—the actual vector, not its components in some frame (for further details, see Penrose 1968b, Penrose & Rindler 1984).

Since from (2.1) we obtain

$$g_{ab}V^aV^b = 2\det(V^{AA'}) = \epsilon_{AB}\epsilon_{A'B'}V^{AA'}V^{BB'}, \tag{2.9}$$

we have the fundamental relation

$$g_{ab} = \epsilon_{AB}\epsilon_{A'B'} \tag{2.10}$$

where the spinors ϵ_{AB} and $\epsilon_{A'B'}$ are complex conjugates of one another

$$\epsilon_{A'B'} = \bar{\epsilon}_{A'B'}(=\overline{\epsilon_{AB}}) \tag{2.11}$$

and are *skew* (i.e. in $\mathbb{S}^* \wedge \mathbb{S}^*$, $\bar{\mathbb{S}}^* \wedge \bar{\mathbb{S}}^*$, respectively):

$$\epsilon_{BA} = -\epsilon_{AB}, \qquad \epsilon_{B'A'} = -\epsilon_{A'B'} \tag{2.12}$$

with $\epsilon_{01} = 1 = \epsilon_{0'1'}$ in standard coordinates. The spinors ϵ_{AB} and $\epsilon_{A'B'}$ establish isomorphisms between $\mathbb{S}$, $\bar{\mathbb{S}}$ and their duals $\mathbb{S}^*$, $\bar{\mathbb{S}}^*$ by means of the index-lowering operation[3]

$$\begin{aligned} \xi^A &\mapsto \xi_A = \xi^B\epsilon_{BA} \\ \eta^{A'} &\mapsto \eta_{A'} = \eta^{B'}\epsilon_{B'A'} \end{aligned} \tag{2.13}$$

the inverse operation

$$\begin{aligned} \xi_A &\mapsto \xi^A = \epsilon^{AB}\xi_B \\ \eta_{A'} &\mapsto \eta^{A'} = \epsilon^{A'B'}\eta_{B'} \end{aligned} \tag{2.14}$$

serving to define the 'raised' ϵ^{AB}, $\epsilon^{A'B'}$ (elements of $\mathbb{S} \wedge \mathbb{S}$ and $\bar{\mathbb{S}} \wedge \bar{\mathbb{S}}$, respectively). Because of (2.10), the spinor lowering and raising operations are compatible with the more familiar tensor lowering and raising operations, using the (symmetric) metric tensors g_{ab}, g^{ab}.

[3]Repeated abstract indices denote contraction, mirroring the summation convention for numerical indices. This is generated from $\boldsymbol{\gamma}(\boldsymbol{\xi}) = \gamma_A\xi^A = \gamma_B\xi^B = \cdots$, where $\boldsymbol{\gamma} \in \mathbb{S}^*$, $\boldsymbol{\xi} \in \mathbb{S}$.

To describe fields, such as the Maxwell field F_{ab} or the charge-current vector J^a, we need cross-sections of bundles. The primary bundle is the spin-vector bundle (fibre $\S$ over each point). The space of its (say C^∞ smooth) cross-sections is a module $\mathfrak{S}^{\cdot}$ over the C^∞ scalar fields on space-time. Its complex conjugate is $\bar{\mathfrak{S}}^{\cdot}$, its dual $\mathfrak{S}^{\cdot *}$ and conjugate dual $\bar{\mathfrak{S}}^{\cdot *}$. The module of complex world vector fields is

$$\mathfrak{V}^{\cdot} \oplus i\mathfrak{V}^{\cdot} = \mathfrak{S}^{\cdot} \otimes \bar{\mathfrak{S}}^{\cdot} \tag{2.15}$$

(module tensor product), where the real world vector fields $\mathfrak{V}^{\cdot}$ are the Hermitian elements, constituting a module over the real C^∞ scalar fields.

We are particularly concerned with spinor fields which are 'pointwise irreducible' in the sense of being symmetric in all unprimed and in all primed indices. For convenience we take these to be *lower* indices; but because of the isomorphisms (2.13), (2.14) this is no effective restriction. Thus we are concerned with the modules

$$\mathfrak{S}_{p,q} = \underbrace{(\mathfrak{S}^{\cdot *} \odot \ldots \odot \mathfrak{S}^{\cdot *})}_{p} \otimes \underbrace{(\bar{\mathfrak{S}}^{\cdot *} \odot \ldots \odot \bar{\mathfrak{S}}^{\cdot *})}_{q} \tag{2.16}$$

($\odot$ meaning symmetric tensor product.) having elements

$$\phi_{A \ldots CD' \ldots F'} = \phi_{(A \ldots C)(D' \ldots F')}. \tag{2.17}$$

Consider, now, the Maxwell tensor F_{ab}, which belongs to $\mathfrak{V}^{\cdot *} \wedge \mathfrak{V}^{\cdot *}$. We can write

$$\begin{aligned} F_{ab} &= F_{AA'BB'} \\ &= \phi_{AB}\epsilon_{A'B'} + \epsilon_{AB}\bar{\phi}_{A'B'} \end{aligned} \tag{2.18}$$

where the second line arises from splitting $F_{AA'BB'}$ into the piece symmetric in AB, and therefore skew in $A'B'$, and the piece skew in AB and therefore symmetric in $A'B'$. (Because of the 2-dimensionality of spin-space, any skew pair of spinor indices can always be 'split off' as an ϵ-spinor.) The spinor ϕ_{AB} is symmetric

$$\phi_{AB} = \phi_{BA} \tag{2.19}$$

and therefore belongs to $\mathfrak{S}_{2,0}$. Its three complex components ϕ_{00}, ϕ_{01}

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

$(=\phi_{10})$ and ϕ_{11} are certain linear combinations of the components of $\boldsymbol{E} - \mathrm{i}\boldsymbol{B}$, so it conveniently codes the electromagnetic field into the spinor framework. Maxwell's equations now take the form

$$\nabla^{A}_{A'}\phi_{AB} = 2\pi J_{BA'}\,. \tag{2.20}$$

Though I have been regarding the operator $\nabla_a = \nabla_{AA'}$ here (and in Section 1) as being the normal flat-space gradient (cf. (1.8)), all the equations, so far, will also apply equally well in the curved space-time of general relativity, ∇_a now denoting standard (Christoffel) covariant derivative.

In the context of quantum mechanics it is significant also to consider *complex* Maxwell fields F_{ab} since these can be used as the wave functions of single photons. In place of (2.18) we now have

$$F_{ab} = \phi_{AB}\epsilon_{A'B'} + \epsilon_{AB}\tilde{\phi}_{A'B'} \tag{2.21}$$

with $\tilde{\phi}_{A'B'}$ a symmetric spinor which is independent of ϕ_{AB}. The two terms on the right-hand side of (2.21) are, respectively, the anti-self-dual part

$$^{-}F_{ab} = \phi_{AB}\epsilon_{A'B'} \tag{2.22}$$

and self-dual part

$$^{+}F_{ab} = \epsilon_{AB}\tilde{\phi}_{A'B'} \tag{2.23}$$

of F_{ab}, satisfying

$$\tfrac{1}{2}\mathrm{e}_{abcd}{}^{\pm}F^{cd} = \pm\mathrm{i}^{\pm}F_{ab}\,, \tag{2.24}$$

i.e., in terms of the corresponding 2-form notation,

$$^{+}\boldsymbol{F}^{*} = \mathrm{i}\,{}^{+}\boldsymbol{F}, \qquad {}^{-}\boldsymbol{F}^{*} = -\mathrm{i}\,{}^{-}\boldsymbol{F}\,. \tag{2.25}$$

Let us now restrict attention to flat (Minkowski) space time $\mathscr{M}$. For the wave-function of a free photon, in place of the classical reality condition on F_{ab}, which would make ϕ_{AB} and $\tilde{\phi}_{A'B'}$ mutually complex conjugate (and therefore $^{-}\boldsymbol{F}$ and $^{+}\boldsymbol{F}$ mutually complex conjugate), we have a *positive frequency* condition on each of the two independent

spinor fields, ϕ_{AB} in $\mathfrak{S}_{2,0}$ and $\tilde{\phi}_{A'B'}$ in $\mathfrak{S}_{0,2}$. This condition can be stated in various ways and means, in effect, that, in terms of a Fourier decomposition of the fields, only Fourier components of the form $e^{-i\omega t}$ appear, with $\omega > 0$, and not those of the form $e^{i\omega t}$. The source-free Maxwell equations are to hold for each of ϕ_{AB} and $\tilde{\phi}_{A'B'}$ separately:

$$\nabla^{A}_{A'}\phi_{AB} = 0\,, \tag{2.26}$$

$$\nabla^{A'}_{A}\tilde{\phi}_{A'B'} = 0\,. \tag{2.27}$$

The two fields ϕ_{AB} and $\tilde{\phi}_{A'B'}$ describe, respectively, the *left-handed* and *right-handed* circularly polarized parts of the photon wave-function (cf. Penrose & MacCallum 1973). For the left-handed part the spin is anti-parallel to the momentum (negative helicity) and for the right-handed part it is parallel to it (positive helicity). In each case, the absolute magnitude of the spin is unity (in units of $\hbar$). Thus, we say that ϕ_{AB}, subject to (2.26), describes a helicity -1 massless particle while $\tilde{\phi}_{A'B'}$, subject to (2.27) describes a helicity $+1$ massless particle.

There is a ready generalization of this to arbitrary integral or half-integral helicity. If $s = \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ we give the wave function for a helicity s massless particle as a (non-zero) element of $\mathfrak{S}_{0,2s}$

$$\underbrace{\tilde{\phi}_{A'\ldots E'}}_{2s} = \tilde{\phi}_{(A'\ldots E')} \tag{2.28}$$

subject to

$$\nabla^{A'}_{A}\tilde{\phi}_{A'B'\ldots E'} = 0\,, \tag{2.29}$$

while if $s = -\frac{1}{2}, -1, -\frac{3}{2}, -2, \ldots$ we give it as an element of $\mathfrak{S}_{-2s,0}$

$$\underbrace{\phi_{A\ldots E}}_{-2s} = \phi_{(A\ldots E)} \tag{2.30}$$

subject to

$$\nabla^{A}_{A'}\phi_{AB\ldots E} = 0\,. \tag{2.31}$$

An immediate consequence of (2.29) and (2.31) is that the field is annihilated by the wave operator

$$\Box = \nabla_a\nabla^a \tag{2.32}$$

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

(which fact serves to justify the term 'massless' field or particle). When $s = 0$ we have no first-order equation like (2.29) or (2.31) and we simply adopt

$$\Box\phi = 0 \tag{2.33}$$

where $0 \neq \phi \in \mathfrak{S}_{0,0}$. In each case the spin $j = |s|$.

When $s = \pm\frac{1}{2}$ these equations have physical interest as the Dirac–Weyl descriptions of a (massless) anti-neutrino or neutrino. When $j = 2$ we can set

$$K_{abcd} = \phi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\tilde{\phi}_{A'B'C'D'} \tag{2.34}$$

and find that K_{abcd} satisfies the symmetry conditions of a (complex) curvature tensor

$$K_{abcd} = K_{[ab][cd]} = K_{cdab}\,, \qquad K_{[abc]d} = 0 \tag{2.35}$$

together with the 'Einstein vacuum equations' (vanishing Ricci tensor)

$$K_{abc}{}^{b} = 0\,. \tag{2.36}$$

The massless field equations (2.29), (2.31) now take the form of the Binachi identity:

$$\nabla_{[a}K_{bc]de} = 0\,. \tag{2.37}$$

It follows that K_{abcd} describes the weak-field curvature in the weak-field limit of Einstein's theory, for a free (i.e. vacuum) gravitational field (cf. Fierz & Pauli 1939, Penrose 1960, Penrose & Rindler 1984).

When $\mathbb{M}$ is replaced by a general curved space-time $\mathcal{M}$, certain unpleasant consistency conditions arise with (2.29), (2.31) which invalidate their direct use except in the cases $s = \pm\frac{1}{2}, \pm 1$ (and it is fairly customary also to replace (2.33) by $(\Box + \frac{1}{6}R)\phi = 0$, for the case $s = 0$). (See Buchdahl 1958, Plebanski 1965, Penrose & Rindler 1984). However, in general relativity proper we *do* have full consistency for a free gravitational field, provided we take K_{abcd} to be (a constant multiple of) the curvature tensor.

3. Twistors

One way in which a twistor may be viewed (in flat space-time $\mathbb{M}$) is as an object supplying a link between massless fields of adjacent helicities (cf. Penrose 1968a, 1975, Penrose & MacCallum 1973, Wess & Zumino

1974). To this end, as may take a twistor to be an element ω^A (or rather, ω_A) of $\mathcal{Z}_{1,0}$ satisfying the *twistor equation*:

$$\nabla^{(A}_{A'}\omega^{B)} = 0\,. \tag{3.1}$$

It is not hard to show that, in $\mathcal{M}$, the solutions of (3.1) constitute a *complex four-dimensional vector space*. I denote this space by $\mathbb{T}$ and refer to it as *twistor space*. The significance of (3.1), for our present purposes, is that if $\phi_{A\ldots E}$ satisfies (2.30) and (2.31), for helicity s (<0), then

$$\phi_{A\ldots DE}\omega^E \tag{3.2}$$

satisfies the corresponding equation for helicity $s+\frac{1}{2}$. Indeed, one readily sees, on expanding out

$$\nabla^A_{A'}(\phi_{AB\ldots DE}\omega^E) = 0 \tag{3.3}$$

that (3.1) is *precisely* the equation needed to ensure that (3.2) maps helicity s massless fields to helicity $s+\frac{1}{2}$ massless fields ($s<0$). The restriction to $s<0$ can also be removed if, for $s\geqslant 0$, we replace (3.2) by

$$\omega^F\nabla_{FF'}\tilde{\phi}_{A'\ldots E'} + (s+\tfrac{1}{2})\tilde{\phi}_{(A'\ldots E'}\nabla_{F')F}\omega^F \tag{3.4}$$

but (3.4) will not concern us here.

This spin-lowering property of (3.2), together with the four-dimensionality of $\mathbb{T}$ and certain relations between $\mathbb{T}$ and angular momentum structure that I shall present shortly, is essentially all that will be needed of twistor theory here. It may be mentioned in passing, however, that there is a great deal more that can be done with twistors; there are many applications to Maxwell theory and to the massless fields of other spins. Indeed, there is a point of view according to which one even seeks to use twistor theory to provide an *alternative framework for physics* to that provided by the normal space-time descriptions. According to this view, points of $\mathcal{M}$ are re-interpreted in terms of two-dimensional linear subspaces of $\mathbb{T}$. Massless fields (including that of Maxwell) find a surprisingly neat and effective transcription into the holomorphic sheaf cohomology of regions in $\mathbb{T}$. Moreover certain non-linear fields, notably the Einstein vacuum and Yang–Mills fields, in the (anti-)self-dual case, can also be transcribed into twistor terms in remarkably natural-looking

ways. In the absence of a more complete transcription which, among other things, would allow for the removal of this (anti-) self-dual restriction for these nonlinear fields, twistor theory remains a long way from achieving its more grandiose aims. But in more limited ways, a fair amount has already been achieved. (For more details, see Penrose 1967, 1968a, 1975, 1976, Ward 1977, Atiyah & Ward 1977, Atiyah, Hitchin, Manin & Drinfeld 1978, Hitchin 1979, Tod & Ward 1979, Wells 1979, Lerner & Sommers 1979, Hughston & Ward 1979, Penrose & Ward 1980, Hughston 1979, Eastwood, Penrose & Wells 1981, Penrose & Rindler 1984.)

It is of some interest to note that the twistor equation (3.1) is, in a certain sense, 'complementary' to the Dirac–Weyl neutrino equation

$$\nabla^{A}_{A'}\phi_{A} = 0\,. \tag{3.5}$$

For considering the full 'gradient' of ϕ_B, we can split is as follows:

$$\nabla_{a}\phi_{B} = \nabla_{A'[A}\phi_{B]} + \nabla_{A'(A}\phi_{B)} = -\tfrac{1}{2}\epsilon_{AB}\nabla^{C}_{A'}\phi_{C} + \nabla_{A'(A}\phi_{B)}\,. \tag{3.6}$$

Ignoring the ϵ_{AB}, we see that the first term of the R.H.S. of (3.6) belongs to $\mathfrak{S}_{0,1}$ and is the L.H.S. of (3.5), whereas the second term of the R.H.S. of (3.6) belongs to $\mathfrak{S}_{3,1}$ and is the (index-lowered version of the) L.H.S. of (3.1). Thus, according to (3.6) we have a map

$$\nabla : \mathfrak{S}_{1,0} \to \mathfrak{S}_{0,1} \oplus \mathfrak{S}_{3,1}\,. \tag{3.7}$$

When the image lies entirely in the second factor we have a solution of the Dirac–Weyl equation; when it lies entirely in the first, we have a solution of the twistor equation.

Similarly, but somewhat more generally, the action of the gradient on $\mathfrak{S}_{p,0}$ $(p > 0)$ splits as follows:

$$\nabla : \mathfrak{S}_{p,0} \to \mathfrak{S}_{p-1,1} \oplus \mathfrak{S}_{p+1,1}\,. \tag{3.8}$$

When the image lies entirely in the second factor we have a solution of the massless free-field equation for helicity $-\frac{1}{2}p$; when entirely in the first, we have a solution of the *symmetric p-index twistor equation*:

$$\nabla^{(A}_{A'}\omega^{B\ldots E)} = 0, \qquad \omega^{\overbrace{B\ldots E}^{p}} = \omega^{(B\ldots E)}\,. \tag{3.9}$$

It turns out (cf. Penrose 1967) that, in $\mathbb{M}$, the solutions of (3.9) constitute a $\frac{1}{6}(p+1)(p+2)(p+3)$-dimensional complex vector space, which may be identified with

$$\underbrace{\mathbb{T}\odot\ldots\odot\mathbb{T}}_{p}. \tag{3.10}$$

Moreover, we can use a solution of (3.9) directly to lower the spin by $\frac{1}{2}p$ units by contracting it directly into a massless free field of helicity $\leqslant -\frac{1}{2}p\,s \leqslant -\frac{1}{2}p$, to form

$$\phi_{A\ldots CD\ldots F}\omega^{D\ldots F}. \tag{3.11}$$

The action of ∇ on a general $\mathfrak{S}_{p,q}(p,q>0)$ is somewhat more complicated

$$\nabla : \mathfrak{S}_{p,q} \to \mathfrak{S}_{p-1,q-1} \oplus \mathfrak{S}_{p-1,q+1} \oplus \mathfrak{S}_{p+1,q-1} \oplus \mathfrak{S}_{p+1,q+1}. \tag{3.12}$$

When the image is entirely in the last factor, it turns out we have a helicity $\frac{1}{2}(p-q)$ massless free field differentiated $\max(p,q)$ times, and when the image *avoids* the last factor, the solutions constitute a $\frac{1}{12}(p+1)(p+2)(q+1)(q+2)(p+q+3)$-dimensional complex vector space which may be identified with *trace-free part of*

$$\underbrace{((\mathbb{T}\odot\ldots\odot\mathbb{T})}_{p}\otimes\underbrace{(\bar{\mathbb{T}}\odot\ldots\odot\bar{\mathbb{T}}))}_{q}. \tag{3.13}$$

Here the complex conjugate $\bar{\mathbb{T}}$, of $\mathbb{T}$, may be identified with its dual $\mathbb{T}^*$

$$\bar{\mathbb{T}} = \mathbb{T}^* \tag{3.14}$$

because of the existence of a canonical Hermitian scalar product

$$\mathbb{T}\otimes\bar{\mathbb{T}}\to\mathbb{C} \tag{3.15}$$

defined by

$$\tfrac{1}{2}\mathrm{i}(\mu^{A'}\nabla_{AA'}\omega^{A} - \omega^{A}\nabla_{AA'}\mu^{A'}), \tag{3.16}$$

this being constant over $\mathscr{S}$ provided ω^A satisfies the twistor equation

(3.1) and $\mu^{A'}$ satisfies the conjugate twistor equation

$$\nabla^{(A'}_{A}\mu^{B')} = 0\,. \tag{3.17}$$

We may regard (3.8) as a particular case of (3.12) if we allow $q = 0$ and define $\mathfrak{S}_{r,-1} = \{0\}$ (zero-dimensional vector space); similarly, if we allow $p = 0$ and define $\mathfrak{S}_{-1,r} = \{0\}$ we obtain the fact that the solution space of the conjugate version of (3.9)

$$\nabla^{(A'}_{A}\mu^{B'\ldots E')} = 0, \qquad \mu^{\overbrace{B'\ldots E'}^{q}} = \mu^{(B'\ldots E')} \tag{3.18}$$

may be identified with the dual of (3.10) (with q in place of p). In fact the duality between solutions of (3.18) with those of (3.9) (taking $q = p$) may be expressed directly, the required scalar product, generalizing (3.16), being

$$\begin{aligned}\left(\frac{\mathrm{i}}{2}\right)^{p}\Big\{&\mu^{A'B'\ldots E'}\nabla_{AA'}\nabla_{BB'}\ldots\nabla_{EE'}\omega^{AB\ldots E}\\ &- p(\nabla_{AA'}\mu^{A'B'\ldots E'})\nabla_{BB'}\ldots\nabla_{EE'}\omega^{AB\ldots E}\\ &+\frac{p(p-1)}{2}(\ldots)-\cdots+(-1)^{p}(\nabla_{AA'}\nabla_{BB'}\ldots\nabla_{EE'}\mu^{A'B'\ldots E})\omega^{AB\ldots E}\Big\}\,,\end{aligned} \tag{3.19}$$

which is constant over $\mathbb{M}$.

The particular case of (3.13) in which $p = q = 1$ has a special interest, since this is the solution space, for vectors $k^a = k^{AA'}$, of the equation

$$\nabla^{(A}_{(A'}k^{B)}_{B')} = 0 \tag{3.20}$$

which may be rewritten

$$\nabla_{(a}k_{b)} = \tfrac{1}{4}g_{ab}\nabla_{c}k^{c}\,. \tag{3.21}$$

Equation (3.21) will be recognized as the *conformal Killing equation*, whose real solutions are the generators of (local) conformal motions of $\mathbb{M}$. The conformal Killing vectors in $\mathbb{M}$ from a 15-dimensional (real) vector space which may be identified with *trace-free Hermitian part of*

$$\mathbb{T}\otimes\bar{\mathbb{T}}\,. \tag{3.22}$$

Particular among the conformal Killing vectors are the *Killing vectors*, which satisfy

$$\nabla_{(a}k_{b)} = 0\,, \tag{3.23}$$

and generate the restricted Poincaré motions (isometries) of $\mathcal{M}$. These form a 10-dimensional real vector space $\mathbb{K}$ which is contained in the real subspace of (3.22) spanned by corresponding elements of

$$\mathbb{S}\otimes\bar{\mathbb{T}}\,, \qquad \mathbb{T}\otimes\bar{\mathbb{S}}\,. \tag{3.24}$$

Here I am regarding $\mathbb{S}$ to be a subspace of $\mathbb{T}$ (and $\bar{\mathbb{S}}$ of $\bar{\mathbb{T}}$) consisting of *constant* ω^A satisfying (3.1) (constant $\mu^{A'}$ satisfying (3.17)). In fact the inclusion $\mathbb{S}\subset\mathbb{T}$ provides us with an exact sequence

$$0\to\mathbb{S}\to\mathbb{T}\to\bar{\mathbb{S}}^*\to 0\,, \tag{3.25}$$

where the reason that the fourth term may be identified as the dual of the complex conjugate of the second lies in the duality (3.14) as expressed explicitly by (3.16). In fact the complex conjugate (3.25) (written in reverse order)

$$0\leftarrow\mathbb{S}^*\leftarrow\bar{\mathbb{T}}\leftarrow\bar{\mathbb{S}}\leftarrow 0 \tag{3.26}$$

may be identified as the *dual* of (3.25) (cf. Penrose & MacCallum 1973). Recall that ϵ_{AB} and ϵ^{AB} establish an isomorphism between $\mathbb{S}^*$ and $\mathbb{S}$. Thus, the second map in (3.26) (and in (3.25)) provides us with

$$\bar{\mathbb{T}}\to\mathbb{S}\hookrightarrow\mathbb{T}\,. \tag{3.27}$$

The map (3.27) applied to one factor of the *symmetric* space

$$\bar{\mathbb{T}}\odot\bar{\mathbb{T}} \tag{3.28}$$

provides us with a subspace of the first space in (3.24), and likewise the complex conjugate operation provides us with a corresponding subspace of the second. The (real) linear span of the sums of corresponding elements from the two spaces is what provides us with the actual space $\mathbb{K}$.

This may seem a little involved, and an (abstract) index notation can be used to help organize things. I use lower case Greek letters for the

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

four-dimensional abstract twistor indices. Thus, Z^α represents an element of $\mathbb{T}$ and W_α an element of $\bar{\mathbb{T}}$. It is convenient sometimes to use the notation

$$Z^\alpha = (\omega^A, \pi_{A'}), \qquad W_\alpha = (\lambda_A, \mu^{A'}) \tag{3.29}$$

where Z^α and W_α respectively represent fields ω^A, satisfying (3.1), and $\mu^{A'}$, satisfying (3.17), and where we find

$$\nabla^A_{A'}\omega^B = -i\epsilon^{AB}\pi_{A'}, \qquad \nabla^{A'}_A\mu^{B'} = i\epsilon^{A'B'}\lambda_A \tag{3.30}$$

define $\pi_{A'}$ and λ_A as *constant* elements of $\mathfrak{S}_{0,1}$ and $\mathfrak{S}_{1,0}$ respectively. In fact (by my adopted abuses of notation) we may regard $\pi_{A'}$ and λ_A to be the elements of $\bar{\mathbb{S}}^*$ and $\mathbb{S}^*$ which are given, respectively, by the third map of (3.25) and the second map of (3.26). The scalar product (3.16) of Z^α and W_α now takes the form

$$Z^\alpha W_\alpha = \omega^A\lambda_A + \pi_{A'}\mu^{A'} \tag{3.31}$$

which, being a field constant over $\mathbb{M}$, may be regarded as an element of $\mathbb{C}$.

For the special case given when W_α is the complex conjugate of Z^α

$$W_\alpha = \bar{Z}_\alpha = (\bar{\pi}_A, \bar{\omega}^{A'}) \tag{3.32}$$

we obtain the 'norm'

$$Z^\alpha\bar{Z}_\alpha = \omega^A\bar{\pi}_A + \pi_{A'}\bar{\omega}^{A'} \in \mathbb{R} \tag{3.33}$$

which is a Hermitian form on $\mathbb{T}$ of signature $(+\,+\,-\,-)$. This defines the isomorphism (3.14).

An element $K^\alpha{}_\beta$ of $\mathbb{K}$, representing a (real) Killing vector k^a, must be Hermitian and trace-free (cf. (3.22))

$$\bar{K}_\alpha{}^\beta = K^\beta{}_\alpha, \qquad K^\alpha{}_\alpha = 0 \tag{3.34}$$

and, according to the above descriptions, of the form

$$K^\alpha{}_\beta = I^{\gamma\delta}\bar{Q}_{\gamma\beta} + Q^{\gamma\alpha}\bar{I}_{\gamma\beta} \tag{3.35}$$

with $Q^{\alpha\beta}$ symmetric

$$Q^{\alpha\beta} = Q^{\beta\alpha}\,, \tag{3.36}$$

where the map (3.27) (which is $(\lambda_A, \mu^{A'}) \to (\lambda^A, 0)$) is achieved by

$$W_a \to I^{\alpha\gamma} W_\gamma\,, \tag{3.37}$$

defining the *infinity twistor*

$$I^{\alpha\beta} = I^{\beta\alpha}\,. \tag{3.38}$$

This twistor is *simple*

$$I^{[\alpha\beta} I^{\rho]\sigma} = 0 \tag{3.39}$$

and *self-dual* in the sense

$$\bar{I}_{\alpha\beta} = \tfrac{1}{2}\epsilon_{\alpha\beta\gamma\delta} I^{\gamma\delta} \tag{3.40}$$

where the alternating twistor

$$\epsilon_{\alpha\beta\gamma\delta} = \epsilon_{[\alpha\beta\gamma\delta]} \tag{3.41}$$

is fixed by (3.41) and a certain natural choice of scaling (in fact determined by (3.40)).

Another object which turns out to have a twistor character is the *angular momentum* structure of a system in special relativity. Taken relative to an origin $o \in \mathbb{M}$, the total angular momentum of a system (assumed finite) is described by an antisymmetric tensor $M^{ab}(o)$ (i.e. element of $\mathbb{V} \wedge \mathbb{V}$). (The six independent components of relativistic angular momentum are the three components of non-relativistic angular momentum together with three components describing the location of the mass centre.) When we pass to a new origin $X \in \mathbb{M}$, whose position vector relative to o is given by x^a, we find that the angular momentum $M^{ab}(X)$, relative to X, is given by

$$M^{ab}(X) = M^{ab}(o) - 2x^{[a}p^{b]} \tag{3.42}$$

where p^a is the total 4-momentum of the system. (Equation (3.42) describes both angular momentum conservation and the uniform motion of the mass-centre as determined by the momentum.) We may think of $M^{ab}(X)$, henceforth written simply as M^{ab}, as a space-time field which, by

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

(3.42), is subject to the 'field equation'

$$\nabla_{(a}M^{*}_{b)c} = 0\,, \tag{3.43}$$

since the momentum p^a is constant, where

$$M^{*}_{ab} = \tfrac{1}{2}\epsilon_{abcd}M^{cd} \tag{3.44}$$

is the Hodge dual of M^{ab}. We can, if desired, rewrite (3.43) as

$$\nabla^a M^{bc} = \tfrac{2}{3}g^{a[b}\nabla_d M^{c]d}\,. \tag{3.45}$$

In spinor terms, we have, as in (2.18)

$$M^{ab} = \bar{\mu}^{AB}\epsilon^{A'B'} + \epsilon^{AB}\mu^{A'B'} \tag{3.46}$$

where

$$\mu^{A'B'} = \mu^{B'A'} \tag{3.47}$$

and, moreover,

$$M^{*ab} = -\mathrm{i}\bar{\mu}^{AB}\epsilon^{A'B'} + \mathrm{i}\epsilon^{AB}\mu^{A'B'}\,, \tag{3.48}$$

so (3.43) translates as

$$\epsilon^{BC}\nabla^{AA'}\mu^{B'C'} + \epsilon^{AC}\nabla^{BB'}\mu^{A'C'} = \text{compl. conj.} \tag{3.49}$$

which, on taking irreducible parts becomes

$$\nabla^{(A'}_{A}\mu^{B'C')} = 0\,, \tag{3.50}$$

$$\nabla^{A}_{B'}\mu^{A'B'} = \nabla^{A'}_{B}\bar{\mu}^{AB} \tag{3.51}$$

and the complex conjugate of (3.50). From (3.50) and (3.18) we see that $\mu^{A'B'}$, which is a spinor form of the self-dual part of M^{ab}, belongs to $\bar{\mathbb{T}}\odot\bar{\mathbb{T}}$. Thus M^{ab} itself may be regarded as a 'real' (or Hermitian) element of

$$(\mathbb{T}\odot\mathbb{T})\oplus(\bar{\mathbb{T}}\odot\bar{\mathbb{T}}) \tag{3.52}$$

where, for consistency with what follows, we take the elements of the two components of this direct sum to be given by $(-2\mathrm{i}\bar{\mu}^{AB}, 2\mathrm{i}\mu^{A'B'})$.

The condition (3.51) tells us that M^{ab} is not a general element of (3.52), but is such that when the operation (3.27) is applied to one factor $\bar{\mathbb{T}}$ of the second component of (3.52) then the same (Hermitian) element of $\mathbb{T} \otimes \bar{\mathbb{T}}$ is obtained as when the complex conjugate operation is applied to one of the $\mathbb{T}$ factors of the first component of (3.52). By the nature of (3.27) and its complex conjugate, this element must clearly lie in $\mathbb{S} \otimes \bar{\mathbb{S}}$. In fact one easily checks that this element is actually the 4-momentum p^a.

In terms of the twistor (abstract) index notation this condition can be stated as follows. Lat $A_{\alpha\beta}$—referred to as the *Kinematic twistor* or *angular momentum twistor*—be the above element of $\bar{\mathbb{T}} \odot \bar{\mathbb{T}}$ defined by $2i\mu^{A'B'}$. Then $\bar{A}^{\alpha\beta}$ will be the corresponding element of $\mathbb{T} \odot \mathbb{T}$ defined by $-2i\mu^{AB}$. The required condition on $A_{\alpha\beta}(=A_{\beta\alpha})$, as stated by (3.51), is

$$A_{\alpha\gamma} I^{\beta\gamma} = \bar{A}^{\beta\gamma} \bar{I}_{\alpha\gamma} . \tag{3.53}$$

We recall (cf. before (3.10)) that $\bar{\mathbb{T}} \odot \bar{\mathbb{T}}$ is ten-complex-dimensional. It is the 'Hermiticity condition' (3.53) that reduces the number of independent components of $A_{\alpha\beta}$ from ten *complex* ones to ten *real* ones—the number needed for specifying the relativistic momentum and angular momentum of a physical system. The space

$$\mathbb{A} \subset \bar{\mathbb{T}} \odot \bar{\mathbb{T}} \tag{3.54}$$

of $A_{\alpha\beta}$ satisfying (3.53) is thus ten-real-dimensional.

Another way of stating this condition on $A_{\alpha\beta}$ is that it be expressible as

$$A_{\alpha\beta} = 2E_{(\alpha}{}^{\gamma} \bar{I}_{\beta)\gamma} \tag{3.55}$$

where $E_\alpha{}^\gamma$ is Hermitian in the twistor sense:

$$\bar{E}^\alpha{}_\beta = E_\beta{}^\alpha . \tag{3.56}$$

That is to say, the element in each component of (3.52) may be obtained from the same Hermitian element of $\mathbb{T} \otimes \bar{\mathbb{T}}$ by one operation (3.27), or its complex conjugate, followed by a symmetrization.

The main significance of all this for us lies in the fact that the ten-real-dimensional Killing vector space $\mathbb{K}$ and the ten-real-dimensional kinematic space $\mathbb{A}$ may be regarded as being *duals* of one another. The (real) scalar product between $K^\alpha{}_\beta$, in $\mathbb{K}$, and $A_{\alpha\beta}$, in $\mathbb{A}$, is

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

$$E_\alpha{}^\beta K^\alpha{}_\beta = E_\alpha{}^\beta I^{\gamma\delta}\bar{Q}_{\gamma\beta} + E_\alpha{}^\beta \bar{I}_{\gamma\beta} Q^{\gamma\alpha} \tag{3.57}$$

$$= \mathrm{Re}(Q^{\alpha\beta} A_{\alpha\beta})$$

where $Q^{\alpha\beta}$ is a 'potential' for $K^\alpha{}_\beta$, as defined by (3.35) (and (3.36)), $E_\alpha{}^\beta$ being a 'potential' for $A_{\alpha\beta}$, as defined by (3.55) (and (3.56)). Although there is freedom of choice for each of $Q^{\alpha\beta}$ and $E_\alpha{}^\beta$, the fact that the L.H.S. of (3.57) is independent of $Q^{\alpha\beta}$, whereas the R.H.S. is independent of $E_\alpha{}^\beta$, implies that the scalar product between $K^\alpha{}_\beta$ and $A_{\alpha\beta}$ as defined by (3.57) must be independent of both freedoms. We note that $\mathbb{K}$ may be viewed either as a subspace of $\mathrm{Herm}(\mathbb{T} \otimes \bar{\mathbb{T}})$ or, via (3.35), as a factor space of $\mathrm{Herm}((\mathbb{T} \odot \mathbb{T}) \oplus (\bar{\mathbb{T}} \odot \mathbb{T}))$. Similarly $\mathrm{Herm}(\mathbb{A} \oplus \bar{\mathbb{A}})$ may be viewed either as a subspace of $\mathrm{Herm}((\bar{\mathbb{T}} \odot \bar{\mathbb{T}}) \oplus (\mathbb{T} \odot \mathbb{T}))$ or, via (3.55), as a factor space of $\mathrm{Herm}(\bar{\mathbb{T}} \otimes \mathbb{T})$. Thus the duality between $\mathbb{K}$ and $\mathbb{A}$ may be viewed in terms of the duality between subspaces and factor spaces. (For more details, see Penrose & Rindler 1984.)

The scalar product (3.57) may be expressed directly in terms of the Killing vector k^a and the angular momentum tensor M^{ab} by

$$\tfrac{1}{2} M^{ab} \nabla_a k_b - \tfrac{1}{3} k_b \nabla_a M^{ab}\,, \tag{3.58}$$

which is constant over $\mathbb{M}$ by (3.23) and (3.45). The equivalence to (3.57) may be checked using (3.19), (3.46), etc. When k^a is a time-translation, then (3.58) provides the *energy* of the system associated with that time-direction; likewise, when k^a is a space-translation, (3.58) provides the relevant *momentum* component; when k^a is a rotation, then (3.58) provides the associated *angular momentum* component.

Of particular note, also, is the following twistor expression for the squared rest-mass $m^2 = p_a p^a$:

$$m^2 = -\tfrac{1}{2} A_{\alpha\beta} \bar{A}^{\alpha\beta} \tag{3.59}$$

(cf. Penrose & MacCallum 1973).

We are now in a position to express the analogue of (1.19) for the weak-field limit of Einstein's general relativity. Consider a 2-surface S, as in Fig. 2, which surrounds (i.e. links) some source tube where, for simplicity, we assume that S lies entirely in the source-free region. Then, in the neighbourhood of S, we have a K_{abcd} satisfying (2.35), (2.36), (2.37), i.e. a ϕ_{ABCD} satisfying the massless freefield equation (2.31) for spin 2. For each choice of $\omega^{AB} (= \omega^{BA})$ satisfying (3.9), i.e. for each element of $\mathbb{T} \odot \mathbb{T}$,

we can construct

$$\phi_{AB} = \phi_{ABCD}\omega^{CD} \tag{3.60}$$

and then, via (2.21), substitute this into (1.19) to obtain a 'charge' value for ϕ_{AB}, corresponding to that choice of ω^{AB}. The family of such 'charges' thus constitutes an object $A_{\alpha\beta}(S)$ which belongs to the dual space to $\mathbb{T} \odot \mathbb{T}$, i.e. to $\bar{\mathbb{T}} \odot \bar{\mathbb{T}}$. (It is perhaps more correct to think of this object as $(A_{\alpha\beta}(S), \bar{A}^{\alpha\beta}(S))$ belonging to the dual Herm $\{(\bar{\mathbb{T}} \odot \bar{\mathbb{T}}) \oplus (\mathbb{T} \odot \mathbb{T})\}$ of Herm$\{(\mathbb{T} \odot \mathbb{T}) \oplus (\bar{\mathbb{T}} \odot \bar{\mathbb{T}})\}$.) We find that $A_{\alpha\beta}(S)$ indeed satisfies (3.53) and describes the *total angular momentum and momentum of the sources surrounded by S* (cf. Penrose 1982).

4. General Relativity

The 'twistor' method described at the end of the last section, for obtaining energy-momentum and angular momentum in the weak-field limit of general relativity appears to have been put forward first in Penrose (1968a), although a different but equivalent procedure had been given much earlier by Sachs and Bergmann (1958). Streubel (1978) suggested a way of adapting the twistor method to full general relativity, so as to provide a definition of angular momentum at null infinity $\mathscr{I}^+$, in an asymptotically flat space-time $\mathscr{M}$. There are, however, certain definite objections to Streubel's particular approach. Indeed, it seems that the most natural procedure in curved space-time is not to try to define objects 'ω^{AB}' directly, as would be required in a curved-space analogue of (3.60), but to think of the spin as being, in effect, lowered in *two steps* $2 \to 3/2 \to 1$ using two choices of an 'ω^{A}' subject to an analogue of the original twistor equation (3.1)

$$\nabla^{(A}_{A'}\omega^{B)} = 0 \tag{4.1}$$

rather than of

$$\nabla^{(A}_{A'}\omega^{BC)} = 0, \qquad \omega^{[BC]} = 0 \tag{4.2}$$

(cf. (3.9)).

The ∇-operator in (4.1) and (4.2) is now to be the (standard Christoffel) *covariant derivative* operator for the space $\mathscr{M}$, with its curved (pseudo-)

metric g_{ab}. The commutation of two such operators provides us with the Riemann curvature tensor

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) V^d = R_{abc}{}^d V^c , \tag{4.3}$$

subject to

$$R_{abcd} = R_{[ab][cd]} = R_{cdab}, \qquad R_{[abc]d} = 0 , \tag{4.4}$$

as with K_{abcd} in (2.35), and

$$\nabla_{[a} R_{bc]de} = 0 \tag{4.5}$$

as in (2.37). Einstein's equation

$$R_{abc}{}^b - \tfrac{1}{2} g_{ac} R_{bd}{}^{bd} = -8\pi G T_{ac} \tag{4.6}$$

for energy-momentum tensor $T_{ab} = T_{ba}$ reduces, in vacuum (i.e. $T_{ab} = 0$), to

$$R_{abc}{}^b = 0 \tag{4.7}$$

as in (2.36). Thus, in vacuum, we have an expression exactly analogous to (2.34)

$$R_{abcd} = \Psi_{ABCD} \epsilon_{A'B'} \epsilon_{C'D'} + \epsilon_{AB} \epsilon_{CD} \bar{\Psi}_{A'B'C'D'} \tag{4.8}$$

where

$$\nabla^A_{A'} \Psi_{ABCD} = 0, \qquad \Psi_{ABCD} = \Psi_{(ABCD)} \tag{4.9}$$

as in (2.31), (2.30). (See Penrose & Rindler 1984 for more details).

In $\mathcal{M}$, the equations (4.1), (4.2) are subject to severe consistency conditions (obtained by differentiating further, symmetrizing and contracting), notably

$$\Psi_{ABCD} \omega^A = 0 \quad \text{and} \quad \Psi_{A(BCD} \omega^A{}_{E)} = 0 \tag{4.10}$$

(cf. Penrose & Rindler 1983), which, in the general case, have no solutions other than $\omega^A = 0$, $\omega^{AB} = 0$. Thus, we cannot directly use the

analogue

$$\phi_{AB} = \Psi_{ABCD}\omega^{CD} \tag{4.11}$$

of (3.60) to obtain a spin 1 massless field satisfying $\nabla^{A}_{A'}\phi_{AB} = 0$. Indeed, it is hardly surprising that we cannot. For if we could, we should be able to define non-trivial integrals like (1.19), with the property that, as in the case of Maxwell theory, the value of the integral is unchanged as the closed 2-surface is moved continuously through a source-free region. Such an integral would describe an energy-momentum or angular momentum component to which no contribution would arise from the free gravitational field itself. However, we know from previous work, most notably that of Bondi (1960) and his co-workers (Bondi, van der Burg & Metzner 1962, Sachs 1962), that free gravitational waves do in fact possess energy-momentum (albeit in a non-local way) and, by implication, angular momentum as well. Thus a non-trivial integral such as that suggested above, to which no contribution from vacuum curvature can arise, is not to be expected in general relativity, and the lack of non-trivial solutions to (4.1) and (4.2) is consistent with this.

If all that we ask for is a *definition* of the energy-momentum and angular momentum components, rather than such an unreasonably strong form of conservation law, then we are on safer ground. All that we require is an appropriate definition of ω^{A} or ω^{AB} *on the 2-surface S itself*, and for this purpose it is natural to ask only that the restriction of (4.1) or (4.2) to S hold. Equations (4.1) and (4.2) are, after all, highly over-determined (six complex equations for two complex components in (4.1), and eight complex equations for three complex components in (4.2)), which is the essential reason that these equations fail in a general $\mathscr{M}$. We find, however, that in the case of (4.2) there are only *two* completely tangential equations for the *three* independent components of ω^{AB}, so the problem is now the other way about: the tangential part of (4.2) is insufficient to restrict ω^{AB} to a finite-dimensional family.

On the other hand, the situation for (4.1) is much happier. Choosing a local normalized spinor basis (o^{A}, ι^{A})

$$o_{A}\iota^{A} = 1 \tag{4.12}$$

at points of S, so that the null vectors

$$l^{a} = o^{A}\bar{o}^{A'} \text{ and } n^{a} = \iota^{A}\bar{\iota}^{A'} \tag{4.13}$$

are orthogonal to S everywhere, where now it is required that S be everywhere *spacelike*, we find the *two* tangential equations

$$\check{\delta}'\omega^0 = \sigma'\omega^1, \qquad \check{\delta}\omega^1 = \sigma\omega^0 \tag{4.14}$$

for the *two* components

$$\omega^0 = \omega_1 = \omega_A \iota^A, \quad \omega^1 = -\omega_0 = -\omega_A o^A \tag{4.15}$$

of ω^A. Here, the notation of Geroch, Held & Penrose (1973) is being used, the operators $\check{\delta}'$ and $\check{\delta}$ reducing to the particular form

$$\check{\delta}' = \nabla_{10'} - \beta', \qquad \check{\delta}_{01'} - \beta \tag{4.16}$$

in (4.14), where

$$\beta' = -o^A \nabla_{10'} \iota_A, \qquad \beta = \iota^A \nabla_{01'} o_A \tag{4.17}$$

and also where

$$\sigma' = -\iota^A \nabla_{10'} \iota_A, \qquad \sigma = o^A \nabla_{01'} o_A \,. \tag{4.18}$$

In addition to being space-like, S is now required to have the *topology of a 2-sphere* S^2. Then one can show, using the Atiyah–Singer Index Theorem, that (4.14) always has at least four linearly independent solutions and, in the general case, has *exactly four* independent solutions. (It seems quite possible that the number is always exactly four.) I denote the solution space of (4.14) by

$$\mathbb{T}(S) \tag{4.19}$$

and assume that this space is, indeed, a four-complex-dimensional vector space.

The object ω^{AB}, at S, can now be defined as an element of

$$\mathbb{T}(S) \odot \mathbb{T}(S) \tag{4.20}$$

which amounts to saying that ω^{AB} is a linear combination of expressions like

$$\underset{1}{\omega^A}\,\underset{2}{\omega^B} \tag{4.21}$$

where each of $\underset{1}{\omega}^A$, $\underset{2}{\omega}^B$ satisfies (4.14). It seems that, in the general case, the elements of (4.20) cannot themselves be directly determined by a differential equation. Thus, instead of (4.11), the expressions

$$\phi_{AB} = \Psi_{ABCD}\underset{1}{\omega}^C\underset{2}{\omega}^D \tag{4.22}$$

can be used to define the required 'gravitational charges' surrounded by S (which, in effect, amounts to lowering the spin in two steps, one using $\underset{1}{\omega}^A$ and the other using $\underset{2}{\omega}^A$.

Explicitly we find

$$A_{\alpha\beta}(S)\underset{1}{Z}^\alpha\underset{2}{Z}^\beta = -\frac{\mathrm{i}}{4\pi G}\oint_S\{\Psi_1\underset{1}{\omega}^0\underset{2}{\omega}^0 + \Psi_2(\underset{1}{\omega}^0\underset{2}{\omega}^1 + \underset{1}{\omega}^1\underset{2}{\omega}^0) + \Psi_3\underset{1}{\omega}^1\underset{2}{\omega}^1\}\mathcal{S} \tag{4.23}$$

where $\underset{i}{Z}^\alpha$ stands for the solution $(\underset{i}{\omega}^0, \underset{i}{\omega}^1)$ of (4.14) ($i = 1, 2$), where $\mathcal{S}$ is the surface-area 2-form of S and where

$$\Psi_3 = \Psi_{0111}, \qquad \Psi_2 = \Psi_{0011}, \qquad \Psi_1 = \Psi_{0001}, \tag{4.24}$$

these components being taken in the basis (o^A, ι^A), as in (4.15) for lower spinor indices. Equation (4.23) serves to define the angular momentum twistor $A_{\alpha\beta}(S)$, as an element of

$$\mathbb{T}^*(S)\odot\mathbb{T}^*(S), \tag{4.25}$$

for the total 'source' surrounded by S, where S is a spacelike topological 2-sphere lying entirely in the vacuum region.

There is also a generalization of (4.23) which allows us to remove the restriction that S lie only in the vacuum region. For this, we need only replace Ψ_3, Ψ_2 and Ψ_1, in (4.23), as follows:

$$\begin{aligned}\Psi_1 &\to R_{abcd}l^a m^b \bar{m}^c m^d \\ \Psi_2 &\to \tfrac{1}{2}R_{abcd}(l^a n^b + \bar{m}^a m^b)\bar{m}^c m^d \\ \Psi_3 &\to R_{abcd}\bar{m}^a n^b \bar{m}^c m^d\end{aligned} \tag{4.26}$$

 Integrals for general-relativistic sources: A development from Maxwell's electromagnetic theory. *The sesquicentennial symposium* (1981). Eds Melvyn S. Beger and J. C. Maxwell, 211–43. 1984

where l^a, n^a are as in (4.13) and

$$m^a = o^A \bar{\iota}^{A'}, \qquad \bar{m}^a = \iota^A \bar{o}^{A'} . \tag{4.27}$$

In order to identify the space (4.25) with the complex conjugate of (4.20), which is necessary if, say, the rest-mass $m(S)$ is to be computed according to (3.59)

$$(m(S))^2 = -\tfrac{1}{2} A_{\alpha\beta}(S) \bar{A}^{\alpha\beta}(S) \tag{4.28}$$

we appear to need an analogue of the identification (3.14)

$$\bar{\mathbb{T}}(S) = \mathbb{T}^*(S) , \tag{4.29}$$

i.e. of the Hermitian scalar product (3.15)

$$\mathbb{T}(S) \otimes \bar{\mathbb{T}}(S) \to \mathbb{C} . \tag{4.30}$$

It is not yet clear what the most appropriate definition for (4.30) should be, but the simplest suggestion seems to be to define

$$\pi_{0'} = \mathrm{i}\check{\delta}'\omega^1 - \mathrm{i}\rho\omega^0, \qquad \pi_{1'} = \mathrm{i}\check{\delta}\omega^0 - \mathrm{i}\rho'\omega^1 , \tag{4.31}$$

these being the parts of the first equation (3.30) which are tangential to S, where in (4.31) the operators $\check{\delta}'$ and $\check{\delta}$ *now* specialize to

$$\check{\delta}' = \nabla_{10'} + \beta', \qquad \check{\delta} = \nabla_{01'} + \beta , \tag{4.32}$$

and then define $Z^a \bar{Z}_a$ to be the *average* of (3.33):

$$\omega^0 \bar{\pi}_0 + \omega^1 \bar{\pi}_1 + \pi_{0'} \bar{\omega}^{0'} + \pi_{1'} \bar{\omega}^{1'} , \tag{4.33}$$

over S.

The quantities (4.31) can also be used in an alternative (equivalent) expression to (4.23), which was found by K.P. Tod:

$$A_{\alpha\beta}(S) \underset{1}{Z}^{\alpha} \underset{2}{Z}^{\beta} = -\frac{\mathrm{i}}{4\pi G} \oint_S (\underset{1}{\pi}_{0'} \underset{2}{\pi}_{1'} + \underset{1}{\pi}_{1'} \underset{2}{\pi}_{0'}) \mathscr{S} . \tag{4.34}$$

As yet, no convincing analogue of $I^{\alpha\beta}$, or the map (3.27), has been found. One might try to average

$$\pi_{1'}\pi_{1'} - \pi_{1'}\pi_{1'} \tag{4.35}$$

over S, which is the expression for $\bar{I}_{\alpha\beta}Z^{\alpha}Z^{\beta}$ in $\mathbb{M}$, but this leads to an $I^{\alpha\beta}(S)$ which is not simple (cf. (3.39)) in a general $\mathcal{M}$. However, the similarity between (4.34) and the average of (4.35) is tantalizing, and it provides some hope that an analogue of (3.53) might exist in the general case. Unfortunately, the correct form of this analogue (if such exists) is still missing, so, as yet, there is no general prescription for reducing the ten complex components of $A_{\alpha\beta}(S)$ to ten real ones.

In general the evaluation of $A_{\alpha\beta}(S)$ is difficult, but promising results have been obtained in a number of different situations. In the first place we know that the expression (4.23) (or (4.34)) gives the expected (correct) answer in the limit of weak fields—which, put another way, means that we also get the expected answer in the limit of small surfaces S in regions of matter (provided that a limiting smooth shape for S is obtained). The expression also works well at null infinity, where a suitable analogue of $I^{\alpha\beta}$ does in fact exist and the relation (3.53) is satisfied. For the energy-momentum components we find full agreement with the highly successful Bondi–Sachs definition. For the angular momentum components we find a somewhat different kind of expression from those which had been given previously. In fact, the present method seems to give a more satisfactory result with regard to the way in which the 'origin', about which the angular momentum is defined, is taken into account. These previous methods are in fact inconsistent with some of the expected answers in the weak field limit, whereas the present method is fully consistent with the weak field limit. (See Penrose 1982, Penrose & Rindler 1984 for further details.)

By using a variety of ingenious arguments, Tod has now been able to evaluate $m(S)$, as given by (4.28), in a number of different situations[4]. He finds, for example, that in the standard Friedmann–Robertson–Walker cosmological models, the mass $m(S)$ surrounded by a spherical surface is equal to that which would be enclosed by a sphere in Euclidean space of the same area, immersed in a fluid of the same density as that of the

[4]Most of his results had not yet been achieved when this lecture was delivered.

model. Thus, in particular, for a model with space-sections which are 3-spheres, $m(S)$ increases to a maximum for an 'equatorial' surface and decreases to zero as the surface shrinks to a point again 'around the back'—giving the result that the total mass of a spatially closed universe is zero, as is normally expected. (This last fact is here a general result because the mass $m(S)$ on one side of S is always equal to the mass $m(S)$ on the other side, for a spatially closed universe.) In the case of the Schwarzschild solution, representing a static spherically symmetrical source, Tod finds the remarkable result that, for surfaces S constrained to one $t = \text{const.}$ hypersurface $m(S)$ is constant and equal to the Schwarzschild 'm' whenever S links the source (once) and is equal to zero whenever S does not link the source. He finds the same result (S now necessarily linking the source) when S is constrained instead to $r = \text{const.}$ or to a symmetrical outgoing null hypersurface. Thus it is natural to conjecture that the same result would hold if these constraints are removed. For the Reissner–Nordstrom charged static spherically symmetrical source, he finds that the results differ from those for the Schwarschild solution in that there is the expected contribution to $m(S)$ from the electromagnetic field energy surrounded by S. He has also examined time-symmetric models representing several sources ('black holes') and finds that for surfaces S in the hypersurface of time-symmetry, $m(S)$ depends only on the sources linked by S and not on the detailed shape of S. The function $m(S)$ is not, however, additive; when two sources are surrounded the result is slightly *less* than the sum of the masses for each source separately, by the expected (negative) potential energy contribution.

This is so far very gratifying, but more work is needed in order to see whether the definitions are always completely satisfactory.

Acknowledgements

I am particularly grateful to K.P. Tod for many suggestions, and also to I.M. Singer, M.F. Atiyah and Tsou S.T.

References

[1] Atiyah, M.F., N.J. Hitchin, V.G. Drinfeld and Yu.I. Manin, (1978), "Construction of instantons, *Phys. Lett. A* 65, 185–187.

[2] Atiyah, M.F. and R.S. Ward, (1977), "Instantons and algebraic geometry", *Comm. Math. Phys.* 55, 117–124.

[3] Bondi, H., (1960), "Gravitational waves in general relativity", *Nature 186*, p. 535.
[4] Bondi, H., M.G.J. van ver Burg and A.W.K. Metzner, (1962), "Gravitational waves in general relativity VII. Waves from axi-symmetric isolated systems", *Proc. R. Soc. Lond. A 269*, 21–52.
[5] Buchdahl, H.A., (1958), "On the compatibility of relativistic wave equations for particles of higher spin in the presence of a gravitational field", *Nuovo. Cim. 10*, 96–103.
[6] Eastwood, M.G., R. Penrose and R.O. Wells Jr., (1981), "Cohomology and massless fields", *Comm. Math. Phys. 78*, 305–351.
[7] Fierz, M. and W. Pauli, (1939), "On relativistic wave equations for particles of arbitrary spin in an electromagnetic field", *Proc. R. Soc. Lond. A 173*, 211–232.
[8] Geroch, R., A. Held and R. Penrose, (1973), "A space-time calculus based on pairs of null directions", *J. Math. Phys. 14*, 874–881.
[9] Hitchin, N.J., (1979), "Polygons and gravitons", *Math. Proc. Camb. Phil. Soc. 85*, 465–476.
[10] Hughston, L.P., (1979), "Twistors and particles", *Springer Lecture Notes in Physics 97*, (Berlin; Springer-Verlag).
[11] Hughston, L.P. and R.S. Ward (eds.), (1979), "Advances in twistor theory", *Research Notes in Mathematics 37*, (San Francisco; London; Melbourne, Pitman).
[12] Lerner, D.E. and P.E. Sommers (eds.), (1979), "Complex manifold techniques in theoretical physics", *Research Notes in Mathematics 32*, (San Francisco; London; Melbourne, Pitman).
[13] Maxwell, J.C., (1861), (1862), "The theory of molecular vortices", *Phil. Mag. 21*, 161–175 (Part I); *21*, 281–291, 338–348 (Part II), *23*, 12–24 (Part III); *23*, 85–95 (Part IV).
[14] Minkowski, H., (1908), "Space and Time", in: *The Principles of Relativity*, (A. Einstein, H.A. Lorentz, H. Weyl and H. Minkowski), (Methuen & Co.) republished (New York (1923), Dover).
[15] Penrose, R., (1960), "A spinor approach to general relativity", *Ann. Phys. 10*, (New York) 171–201.
[16] Penrose, R., (1967), "Twistor algebra", *J. Math. Phys. 8*, 345–366.
[17] Penrose, R., (1968a), "Twistor quantisation and curved space-time", *Internat. J. Theor. Phys. 1*, 61–99.
[18] Penrose, R., (1968b), "Structure of space-time", in *Battelle Recontres, 1967 Lectures in Mathematics and Physics*, eds., C.M. DeWitt and J.A. Wheeler (Benjamin, New York).
[19] Penrose, R., (1975), "Twistor theory, its aims and achievements", in: *Quantum Gravity, an Oxford Symposium*, eds.: C.J. Isham, R. Penrose and D.W. Sciama (Clarendon Press, Oxford).
[20] Penrose, R., (1976), "Nonlinear gravitons and curved twistor theory", *Gen. Rel. Grav. 7*, 31–52.
[21] Penrose, R., (1982), "Quasi-local mass and angular momentum in general relativity", *Proc. R. Soc. Lond. A 351*, 53–63.
[22] Penrose, R. and M.A.H. MacCallum, (1973), "Twistor theory: an approach to the quantisation of fields and space-time", *Phys. Repts. 6*, 242–315.
[23] Penrose, R. and W. Rindler, (1984), *Spinors and space-time*, Vol. 1: *Two-spinor Calculus and Relativistic Fields*, Vol. 2: *Spinor and Twistor Methods in Space-time Geometry*, (Cambridge University Press).
[24] Penrose, R. and R.S. Ward, (1980), "Twistors for flat and curved space-time", in: *General Relativity and Gravitation* ed.: A. Held, vol. 2 (Plenum, New York, London).
[25] Plebanski, J., (1965), "The vectorial' optics of fields with arbitrary spin, rest-mass zero", *Acta Phys. Polon. 27*, 361–393.

[26] Sachs, R.K., (1962), "Gravitational waves in general relativity. VIII Waves in asymptotically flat space-time", *Proc. R. Soc. Lond. A 270*, 103–126.
[27] Sachs, R.K. and P.G. Bergmann, (1958), "Structure of particles in linearized gravitational theory", *Phys. Rev. 112*, 674–680.
[28] Streubel, M., (1978), "Conserved quantities for isolated gravitational systems", *Gen. Rel. Grav. 9*, 551–562.
[29] Tod, K.P. and R.S. Ward, (1979), "Self-dual metrics with self-dual Killing vectors" *Proc. R. Soc. Lond. A 368*, 411–427.
[30] Tolstoy, I., (1981), *James Clerk Maxwell, A Biography*, (Canongate, Edinburgh).
[31] Ward, R.S., (1977), "On self-dual gauge fields", *Phys. Lett. A 61*, p. 81.
[32] Wells, R.O. Jr., (1979), "Complex manifolds and mathematical physics", *Bull. Amer. Math. Soc. (N.S.) 1*, 296–336.
[33] Wess, J. and B. Zumino, (1974), "Supergauge transformations in four dimensions", *Nucl Phys. 70*, 39–50.

153
MASS AND ANGULAR MOMENTUM AT THE QUASI-LOCAL LEVEL IN GENERAL RELATIVITY

Asymptotic behaviour of mass and space-time geometry (1983), 23–30. Springer Lecture Notes in Phys., 202, 1984

Some text is missing in the original article on p.[28], between lines 12 and 13.

378 Mass and angular momentum at the quasi-local level in general relativity. *Asymptotic behaviour of mass and space-time geometry* (1983), 23–30. Springer Lecture Notes in Phys., 202, 1984

MASS AND ANGULAR MOMENTUM AT THE QUASI-LOCAL LEVEL IN GENERAL RELATIVITY

by Roger Penrose
Department of Mathematics, Rice University, Houston, Texas
and
Mathematical Institute, Oxford, U.K.

It has been customary to think of the concept of mass-energy, in general relativity, as accurately applying only at the local level (i.e. as a component of the energy-momentum tensor) or else asymptotically (as with the Bondi-Sachs mass or A.D.M. mass). For a space-time which is assumed to be stationary (or if other simplifying assumptions are made) more can be said. But in the general case, the non-local nature of gravitational wave energy or of granvitational potential energy had seemed to argue strongly against the existence of any rigorously defined intermediate measure of mass-energy which could adequately take into account such gravitational contributions within finitely bounded regions. Recently, I have suggested an expression (Penrose 1982, 1983, Penrose and Rindler 1984) which offers hope of circumventing these difficulties. I shall give here a status report of work concerning this expression, without attempting to describe its details.

The expression refers to a spacelike 2-surface S, whose topology is that of 2-sphere, drawn arbitrarily in an arbitrary space-time M. One obtains a complex of quantities which are intended to measure the total energy-momentum and angular momentum (of matter plus gravitation) which is surrounded by (or "threads through", in the 4-dimensional sense) the surface S. These quantities provide the components of a "twistor" $A_{\alpha\beta}(S)$, belonging to the dual of the symmetric tensor product of the 2-surface twistor space $\mathbb{T}(S)$ with itself. The space $\mathbb{T}(S)$ consists of the solution pairs (ω^0,ω^1) of the 2-surface twistor equations:

$$\eth'\omega^0 = \sigma'\omega^1, \quad \eth\omega^1 = \sigma\omega^0 \tag{1}$$

Here the compacted spin-coefficient notation of Geroch, Held and Penrose (1973) is being used (cf. Newman and Penrose 1962, Penrose and Rindler 1984), the spin-frame being chosen with "flagpoles" orthogonal to S, so σ and σ' are the (complex) shears, at S, of the two null hypersurface through S. In the generic case, $\mathbb{T}(S)$ is four-complex-dimensional. It is not known whether there are (physically reasonable) situations in which $\mathbb{T}(S)$ is not four-dimensional. If there are, the

dimension can only be larger than four, and it might be that in any such circumstance $\mathbf{T}(S)$ could be redefined as a suitable four-dimensional subspace of the solution space of (1). Assume $\mathbf{T}(S)$ to have dimension four. Denoting a typical solution (ω^0,ω^1) of (1) by z^α, where α is a four-dimensional abstract (twistor) index, we define $A_{\alpha\beta}(S)$ by

$$A_{\alpha\beta}(S)z^\alpha z^\beta \equiv -\frac{i}{4\pi G}\oint\{(\Psi_1-\Phi_{01})(\omega^0)^2+2(\Psi_2-\Phi_{11}-\Lambda)\omega^0\omega^1 + (\Psi_3-\Phi_{21})(\omega^1)^2\}dS \qquad (2)$$

where dS is the surface-area element on S, the spin-coefficient notation $\Psi.$, $\Phi..$, being employed for curvature quantities in M. (G is Newton's gravitational constant.)

For a finite surface S, no natural way has yet emerged of singlin out those components of $A_{\alpha\beta}(S)$ which should be regarded as referring to energy-momentum, as opposed to angular momentum. (Perhaps, in the general case no such natural procedure exists.) Related to this question is the (also unsolved) problem of finding a Hermiticity property of $A_{\alpha\beta}(S)$ which reduces its potential ten complex components (10 = 4×5/2) down to ten real ones. However at spacelike infinity (in the cases where the magnetic Weyl curvature falls off fast enough) or at null infinity, or in the linear limit of weak gravitational fields, all these problems are satisfactorily resolved. In the weak-field limit, complete agreement is obtained with the work of Sachs and Bergman (1958) (cf. also Penrose 1968). At null infinity, agreement is obtained with the Bondi-Sachs expression for energy-momentum (cf. Bondi, van der Burg and Metzner 1962, Sachs 1962), while for angular momentum just six independent real components are obtained, precisely agreeing with the six that also arise in the weak-field limit and which are associated with Lorentz-rotational symmetries of Minkowski space. This occurs whether or not the cut (i.e. cross-section) S of $\mathcal{J}^+$ is a "good cut" (the case $\sigma = 0$), so the ambiguities which have arisen with previous definitions of angular momentum (i.e. their inability to distinguish between true angular momentum components and the super-angular momenta which arise from general B.M.S. generators, cf. e.g. Winicour 1980, Geroch and Winicour 1981) do <u>not</u> arise here. At spacelike infinity, full agreement is obtained with the A.D.M. definition of energy-momentum (Arnowitt, Deser and Misner 1961) and (when the magnetic Weyl tensor falls off suitably) with the Ashtekar-Hansen (1978) definition of angular momentum, as has been shown by

Shaw (1983).

For a finite surface S, because of the above-mentioned difficulties in separating the momentum and angular momentum terms in general, the clearest results are obtained when we examine the quasi-local rest-mass m(S), defined by

$$m(S) = \{-\tfrac{1}{2}A_{\alpha\beta}(S)\ \overline{A}^{\alpha\beta}(S)\}^{\frac{1}{2}} \quad . \tag{3}$$

For (3) to be properly defined, a suitable definition for the complex conjugate $\overline{A}^{\alpha\beta}(S)$, of $A_{\alpha\beta}(S)$, is needed, this being an element of the symmetric tensor product of $\mathbf{T}(S)$ with itself. This would be induced once a suitable Hermitian norm (of signature ++--) is provided for $\mathbf{T}(S)$, yielding an anti-linear isomophism between $\mathbf{T}(S)$ and its dual. If M is conformally Minkowskian, the following expression for the norm $Z^{\alpha}\overline{Z}_{\alpha}$, of Z^{α}, can be used:

$$\omega^0\overline{\pi}_0 + \omega^1\overline{\pi}_1 + \pi_0,\overline{\omega}^{0\prime} + \pi_1,\overline{\omega}^{1\prime} \tag{4}$$

where

$$\pi_{0\prime} = i(\eth'\omega^1 - \rho\omega^0),\ \pi_{1\prime} = i(\eth\omega^0 - \rho'\omega^1) \tag{5}$$

(ρ and ρ' measuring the convergences, at S, of the two null hypersurfaces through S). The expressions (4) and (5) come from the standard twistor theory of (conformally) flat space-times (cf. Penrose and MacCallum 1973). It follows from the conformal flatness of M that the expression (4) is in fact constant on S and so can serve as a definition of $Z^{\alpha}\overline{Z}_{\alpha}$ for such space-times. However, in the general case, (4) is not constant and various alternative possibilities suggest themselves. The simplest of these is just to take the average of (4) over S as the definition of $Z^{\alpha}\overline{Z}_{\alpha}$, and I shall refer to this as the average norm. It is not at all clear, as yet, whether the average norm should be adopted or else some other choice made, among the many possibilities. this must be regarded as one of the major unsolved problems of the subject.

One objection that can be levelled at the average norm is that it is not invariant under conformal rescalings of the metric of M (whereas (4) is itself pointwise conformally invariant, so a conformally invariant norm is provided whenever (4) is constant over S). This leads to some difficulties when S becomes a cut of J^+ since the averaging procedure depends upon the choice of conformal factor that is taken to make J^+ finite. It turns out that the rest-mass defined by (3) is actually insensitive to this ambiguity, at J^+. However,

the definition of angular momentum at $\mathcal{J}^+$ does, strictly speaking, require that this ambiguity be resolved. Otherwise one cannot single out the "real origins", about which the angular momentum is to be viewed as being taken. Without an adequate norm, all that one obtains is a complex Minkowski space of possible origins. The locations of the "real" subspace of this complex space is required in order that the total spin of the system be defined. One way to resolve the ambiguity in the norm at $\mathcal{J}^+$ is to use the average norm, where the metric on S is chosen to be that of the standard unit 2-sphere taken to correspond to the direction of the Bondi-Sachs 4-momentum vector as measured at S. This is well defined, but somewhat inelegant.

A related problem is that of finding a suitable flux law for the angular momentum carried by gravitational radiation escaping between two cuts S, S' of $\mathcal{J}^+$. In general there seems to be no clear way of identifying $\mathbb{T}(S)$ with $\mathbb{T}(S')$, the corresponding Minkowski spaces of origins being related by an unknown translation. This problem is considerably more serious in the case of finite surfaces S,S', the unknown translation being replaced by an unknown conformal motion.

Most of the positive results that have been so far achieved have been obtained by Tod (1983). He has, in particular, derived the remarkable theorem that if M is a Schwarzschild space-time (static field of a spherically symmetrical mass or black hole), and if H is any smooth hypersurface in M given by an equation

$$t = f(r) \tag{6}$$

where t and r are the usual time and radial Schwarzschild coordinates (t = const. being an obvious special case), then the mass $m(S)$ given by (3), for a 2-surface lying in the vacuum region of H, depends only upon the linking number of S with the source. If S links the source once, then $m(S)$ is the Schwarzschild mass m; if S does not link the source, then $m(S) = 0$. This leads to the natural conjecture that the same result should hold whether or not S is constrained to a hypersurface of the form (6). It should be pointed out, however, that in the general case the definition of $m(S)$ depends on the choice of norm, whereas when (6) holds it turns out that (4) is constant on S so the norm is unambiguous. Tod has also examined a certain corresponding situation for the Reissner-Nordstrom space-times and finds that there is an (expected) contribution to $m(S)$ from the electromagnetic energy density. It may be remarked that whereas a similar result may be obtained by using the local mass expression of Komar (1959), this yields a contribution from the electromagnetic

field energy which is in error by a factor of 2, as may be verified by examining the limit of weak fields.

Tod has also studied the cases when H is a conformally flat hypersurface of time-symmetry in M. If $H_1, H_2, \ldots, H_n$ are the source regions (or black-hole regions) on H (assumed disjoint and bounded), then the value of $m(S)$, for S lying in the vacuum region in H, depends only upon its linking properties with these various source regions. Thus, one obtains values $m_1, m_2, \ldots, m_n$, where $m(S) = m_j$ whenever S links only the source region H_j (just once), and $m(S) = 0$ when S links no source region. Similarly, one obtains values m_{jk} for $m(S)$ whenever S links just two source regions H_j, H_k. It then turns out that

$$m_{jk} < m_j + m_k \qquad (7)$$

where, in the case of large separations (or weak) fields, the difference between the two sides of (7) is numerically consistent with what one would expect from the presence of a (negative) Newtonian potential energy contribution to the combined sources. A similar result holds when more than two sources are surrounded. (Of course, a strong result like this cannot be expected to hold in a general M, or away from this particular H, since gravitational waves, when present, should contribute a positive mass in the vacuum region.)

A striking feature of the above situation arises from the fact that whereas an inequality holds in (7), the corresponding full quantities $A_{\alpha\beta}(S)$ are <u>additive</u>. One may liken this to the fact that when two energy-momentum 4-vectors are added, their Lorentz lengths (i.e. rest-masses) are generally related only by an inequality. However, in that case the inequality runs in the opposite direction from the one in (7)! The reason for this apparent (and fortunate) discrepancy lies in the fact there is a mixing between energy-momentum and angular momentum components (of the "motion of mass-centre" variety) when one passes from one $A_{\alpha\beta}(S)$ to another. It seems that in the present approach, gravitational potential energy shows up in this curious way.

Another family of examples examined by Tod are the Friedmann-Robertson-Walker cosmological models. The analysis is made easier by the fact that these models are all conformally flat. In particular, Tod finds that for a spherically symmetrical S, the value of $m(S)$ is just what it would have been had S been an ordinary symmetrical 2-sphere, of the same area, in Euclidean 3-space, immersed in a fluid

Asymptotic behaviour of mass and space-time geometry (1983), 23–30. Springer Lecture Notes in Phys., 202, 1984

of the same density as that of the model. As a consequence of this, if S is spherically symmetrical and M is spatially closed (with topology $S^3 \times \mathbb{R}$), $m(S)$ increases monotonically as S increases in size from a tiny surface, until it becomes an "equatorial" 2-sphere. Then, as S shrinks back down on the "other side" of the universe model, $m(S)$ reduces down to zero. Thus, the entire mass (matter plus gravitationally potential energy) of this spatially closed model must be zero, in accordance with some previous expectations. Indeed, this must hold in <u>any</u> spatially closed M, and moreover, the mass $m(S)$ on one side of any S must always be equal to that on the other side, as follows from the symmetry of the expression (2). It would be nice to have a good physical reason for why this should be so. bility of being able to find a hypersurface H which can be embedded in some conformally flat space-time, the intrinsic metric of H, and its extrinsic curvature, being unchanged from that in M. In such cases, the local norm (4) turns out to be constant on S. To gain further insights, one would like to be able to proceed under more general circumstances. Tod has also found it helpful to use his alternative expression

$$A_{\alpha\beta}(S)Z^{\alpha}Z^{\beta} \equiv -\frac{i}{2\pi G}\oint \pi_0 \pi_1' \, dS, \qquad (8)$$

which is equivalent to (2), and which is remarkably simple in appearance. There are some relations between the expression (8) and Witten's (1981) expression for mass, as used in his positivity proof. Indeed, a crucial requirement for the reasonableness of the present definition would be a corresponding positivity theorem for $m(S)$. Moreover, the situation is much more critical here. If the quantity in the curly bracket in (3) can ever be negative, then we shall be presented with an <u>imaginary</u> $m(S)$. And we have seen that in various circumstances this quantity can be zero without M being flat. Whatever the future reveals, it should be interesting!

REFERENCES

1. Arnowitt, R., Deser, S. and Misner, C.W. 1961 Phys. Rev. <u>122</u>, 997.
2. Ashtekar, A. and Hansen, R.O. 1978 J. Math. Phys. <u>19</u>, 1542.
3. Bondi, H., van der Burg, M.G.J. and Metzner, A.W.K. 1962 Proc. R. Soc., Lond. A <u>270</u>, 103.
4. Geroch, R., Held, A. and Penrose, R. 1973. J. Math. Phys. <u>14</u>, 874.
5. Geroch, R., and Winicour, J. 1981. J. Math. Phys. <u>22</u>, 803.
6. Komar, A. 1959. Phys. Rev. <u>113</u>, 934.
7. Newman, E.T. and Penrose, R. 1962. J. Math. Phys. <u>3</u>, 566; <u>4</u>, 998.
8. Penrose, R. 1968. Int. J. Theor. Phys. <u>1</u>, 61.

9. Penrose, R. 1982. Proc. R. Soc. Lond. A 381, 53.
10. Penrose, R. 1983 in J.C. Maxwell, the Sesquicentennial Symposium (ed. M.S. Berger). Elsevier Science Publ. B.V.
11. Penrose, R. and MacCallum, M.A.H. 1973. Phys. Rep. 6C, no. 4, 241.
12. Penrose, R. and Rindler, W. 1984. Spinors and Space-time, vols. 1 & 2 (Cambridge Univ. Press).
13. Sachs, R.K. 1962. Proc. R. Soc. Lond. A 354, 379.
14. Sachs, R.K. and Bergman, P.G. 1958, Phys. Rev. 112, 674.
15. Shaw, W.T. 1983. Proc. R. Soc. Lond. A 390, 191.
16. Tod, K.P. 1983. Proc. R. Soc. Lond. A 333, 457.
17. Winicour, J. 1980. In General relativity and gravitation (ed. A. Held), vol. 2. New York: Plenum."
18. Witten, E. 1981. Communs. Math. Phys. 80, 381.

QUESTIONS

1. Q. What conditions on the Cauchy data on a hypersurface in a given spacetime permit one to use Tod's hypersurface twistor method for calculating the quasi-local mass?

   A. Tod shows that $B_{ab} = 0$ and $\nabla_{[a}E_{b]c} = 0$ are sufficient conditions, where B_{ab} is the magnetic part of the Weyl curvature and E_{ab} is the electric part (relative to a hypersurface).

2. Q. For what spacetime models have calculations of angular momentum been done? For Kerr? For Sato?

   A. Only for Kerr, with very simple surfaces.

3. Q. Have any quasi-local mass calculations been done in perturbed Schwarzschild spacetime?

   A. Not yet, but work on this is presently being done by one of my students.

4. Q. Have any quasi-local mass calculations been done using non-spherical 2-surfaces in a Robertson-Walker spacetime?

   A. Yes. Robertson-Walker spacetimes are conformally flat, so this is not difficult.

5. Q. Why is it difficult to separate off angular momentum?

   A. Because the hermiticity condition is needed here, and no one has figured out what this should be, even in the simple cases. There hasn't been much work on this yet.

6. Q. Is there a twistorial formulation for quasi-local charge in solutions of the Yang-Mills equations?

   A. Tod is working on this. It has the same problems assoicated with it as does the calculation of angular momentum in general relativistic spacetimes.

7. Q. In the general nonflat case, are Killing vector fields very helpful in calculating quasi-local mass?

   A. When there is lots of symmetry (e.g., in the Schwarzschild, the Robertson-Walker, and the Kerr spacetimes) the spacetime

is simple enough to premit quasi-local mass calculations. But individual Killing vector fields have not been successfully exploited. There are two problems: a.) Aligning the surfaces with the Killing fields; b.) Accounting for derivatives off the surfaces (which do matter in quasi-local mass calculations).

8. Q. What properties of "mass" are you trying to capture in your formulation of quasi-local mass? There are at least 15 of them which you might consider.

A. I honestly don't have a complete answer to that question. We want agreement with the weak field limit. We want agreement on J^+ and i_0. We would like positivity.

Physical intuition is important in mathematical-physics. But must be malleable (allowing certain principles to be dropped, if necessary).

154
OBJECTIVE STATE-VECTOR REDUCTION?

Twistor Newsletter **19** (1985)

For quite some while I toyed with this sort of idea for a criterion for a gravitationally induced quantum state reduction. See, for example Chapter 166 and my book *The Emperor's New Mind* (OUP, 1989). I think that it was from discussions with Bob Wald that I finally became persuaded that something different was needed. But only by around 1992, see Chapters 215 and 218, where the idea is to look for a *time-scale* for collapse rather than an absolute on/off criterion, did my views begin to shift to what is roughly my present position. (Lajos Diósi had put forward a rather similar scheme to my later proposal about 8 years earlier: Diósi, L. (1984) Gravitation and quantum mechanical localization of macro-objects *Phys. Lett.* **105A**, 199–202.) See also the updating remarks at the end of Chapter 165.

Objective State-Vector Reduction?

It has been my opinion for a considerable number of years that the resolution of the problem of state-vector reduction in quantum-mechanics should be one of the major tasks of a (successful) quantum gravity theory. This seems to be very much a minority view. Most physicists seem reluctant to monkey with the standard rules of quantum theory (though they are less inhibited about monkeying with the rules of general relativity!). It seems that many physicists do not even accept that there is a problem of state-vector reduction! Here I shall just indicate some of my more recent thoughts on the subject, since there is a vague possibility of a link with twistor theory. Most of what I shall say is spelled out at rather greater length in my contribution to the O.Q.G.D.C. volume: 'Gravity and state-vector reduction' in *Quantum Concepts in Space and Time*, eds. C. J. Isham & R. Penrose (O.U.P., 1985).

I have written at length elsewhere (cf. R.P.: 'singularities and time-asymmetry' in *General Relativity – an Einstein Centenary Survey*, eds. S. W. Hawking & W. Israel (C.U.P., 1979); R.P.: 'Time-asymmetry and quantum gravity' in *Quantum Gravity 2*, eds. C. J. Isham, R. Penrose & D. W. Sciama (O.U.P. 1981)) explaining my views that the 2nd Law of Thermodynamics and cosmological uniformity must arise because of a constraint placed on *initial* space-time singularities (vanishing of Weyl curvature) which should be one implication of a (correct) quantum gravity theory; and that the 'other

side of the coin' is that quantum gravity should be playing an essential role in an objective state-vector reduction. An important ingredient of both these roles for quantum gravity would be a concept of gravitational entropy. The idea, in essence, is that a classical space-time geometry $\mathcal{M}$ would have (perhaps in relation to a spacelike hypersurface $\mathcal{H} \subset \mathcal{M}$) an entropy which measures (the expectation value of) the number of quantum states which contribute to the given classical geometry.

In relation to state-vector reduction, the idea is that a spreading wave function can (always will?) reduce as soon as by doing so, the gravitational entropy goes up by at least as much as the entropy comes down in the wave-function collapse. This is still rather vague, but a rough calculation can be done to see whether these ideas are 'on the right planet'. We anticipate some integral expression involving the Weyl curvature which should have, as two special cases, the Bekenstein-Hawking formula

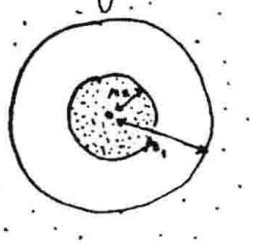
$$\tfrac{1}{4}A \qquad \text{------ (1)}$$

for the entropy of a stationary black hole of surface area A (where Boltzmann's constant, the velocity of light, $\frac{1}{2\pi} \times$ Planck's constant and the gravitational constant are all set equal to unity) and an expression for the 'graviton number' of a weak (dispersed) gravitational wave. For the latter, we envisage an integral over $\mathcal{H}$ of a 'density' of the general form

$$i\int \Psi^+ \iint \overline{\Psi^+} - i\iint \Psi^+ \int \overline{\Psi^+} \qquad \text{... (2)}$$

where Ψ^+ is the positive frequency part of the anti-self-dual linearized Weyl curvature.

Let us imagine the following situation: a nucleus emits a particle in a spherically symmetrical wave function, and this particle is then detected in a cloud chamber or bubble chamber. The detection involves a collapse of the wave function and the idea is that this collapse takes place objectively as soon as the clumping involved in the formation of the droplet or bubble entails an increase of gravitational entropy of at least the amount of the entropy reduction of the wave function collapse. Consider the contraction of a sphere of radius r_1 to one of radius r_2 (where, for simplicity, I provisionally assume that a vacuum is left between them. The effective gravitational entropy 'density' would appear to be something like $\frac{m}{r^2}\cdot\frac{m}{r}$ (since $\Psi \sim \frac{m}{r^3}$) where m is the mass inside r_2 — though it is quite unclear to me as yet how one could rigorously apply these ideas (or what to do with 'positive frequency' when sources are present) — which, when integrated between r_1 and r_2, gives



$$\text{const.} \times m^2 \log\left(\frac{r_1}{r_2}\right). \qquad \ldots (3)$$

This can't be quite right (e.g. for r_2 near the Schwarzschild radius we require $4\pi m^2$, not $m^2 \log(2m)$) but will serve as a rough guide. If we require a gravitational entropy of $\sim 10^{\alpha}$ to reduce the spherical wave to a closely

linear track, we would need a single droplet size of something like a millimeter across — or a string of droplets of somewhat smaller size. In the case of bubbles the result is similar since they can be considered as droplets of negative mass superimposed on a background of uniform density. The m^2 in the formula counts positively ('positive clumping') even when m is negative.

Though this is rather large, by comparison with one's prior expectations as to when systems become 'classical' it is not contradictory and not totally outlandish (i.e. 'on the right planet') — and, in view of all the uncertainties involved, I would call this 'reasonable encouragement' for the idea. (It is the logarithm in ③ that saves things. Gravitational self-energy, rather than entropy, would lead to $(\frac{1}{r_2}-\frac{1}{r_1})m^2$ instead which would be much much too small!)

What has this got to do with twistor theory? Well, perhaps by analogy with the quasi-local mass expression, twistor ideas might turn out to be useful for defining a decent gravitational entropy formula. Gravitational energy is already non-local but 2-surface twistor ideas seem to provide a means of getting at the concept in a quasi-local way.* Gravitational entropy should be 'one stage' less local (in the sense of one more $\int$ in ②. Non-locality seems important for 'reduction' generally. Can twistors help here? ~Roger Penrose

*A comment on the status of quasi-local mass is in order. As of now, there is no evidence against the 'new improved' expression of TN18. But the proof of 'Property K' for Schwarzschild is still incomplete.

CHAPTER 154

155
EMBEDDING 2-SURFACES IN CM

Twistor Newsletter **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

 Embedding 2-surfaces in CM. *Twistor Newsletter* **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.22 **Embedding 2-Surfaces in CM** *by R.Penrose* (**TN** 20, September 1985)

A result of 2-surface twistor theory—apparently of the status of a 'folk-lore theorem'—is that any spacelike 2-surface $\mathcal{S} \subset \mathcal{M}$ for which the space of 2-surface twistors is 4-dimensional (e.g. a generic S^2) can be embedded in complex conformally flat 4-space in such a way that σ and σ' remain unchanged (except for the standard scalings) whereas $\bar{\sigma}$ and $\bar{\sigma}'$ get replaced by new quantities $\tilde{\sigma}$ and $\tilde{\sigma}'$. This embedding (or perhaps immersion) is to be achieved as follows: We have ('natural' immersion)

$$\mathcal{S} \mapsto \mathbf{CM}^{\#}(\mathcal{S}) \tag{1}$$

where the points of $\mathbf{CM}^{\#}(\mathcal{S})$ are the 2-dimensional linear subspaces of $\mathbf{T}^*(\mathcal{S})$ (or, equivalently, of $\mathbf{T}(\mathcal{S})$). To obtain (1), first identify each pair $(p,\ \lambda_A)$—for which $p \in \mathcal{S}$ and λ_A is a spin-covector at p in $\mathcal{M}$—with the element of $\mathbf{T}^*(\mathcal{S})$ which carries the 2-surface twistor $\{\omega^A\}$ to $\omega^A\lambda_A$ (at p). Then p itself gets identified with the linear span of all the λ_A at p.

Now it turns out that the 2-surface twistors on $\mathcal{S}$ can be identified with the restrictions to $\mathcal{S}$ of the solutions of the twistor equation

$$\nabla^{(A}_{A'}\omega^{B)} = 0 \tag{2}$$

in $\mathbf{CM}^{\#}(\mathcal{S})$—which is a (conformally flat) complex manifold, so (2) is meaningful. To see this, note first that the solutions of (2) are precisely the spin-vector fields in $\mathbf{CM}^{\#}(\mathcal{S})$ for which $\omega^A\lambda_A$ is constant over each β-plane in $\mathbf{CM}^{\#}(\mathcal{S})$—where λ_A is such that it is covariantly constant along the β-plane and $\lambda^A\xi^{A'}$ is tangent to it for all $\xi^{A'}$. This follows from the vanishing of

$$\lambda^A\xi^{A'}\nabla_{AA'}(\lambda_B\omega^B) = \xi^{A'}\lambda^A\lambda_B\nabla_{AA'}\omega^B = -\xi^{A'}\lambda_A\lambda_B\nabla^{(A}_{A'}\omega^{B)}. \tag{3}$$

Second, note that the same equation holds at $\mathcal{S}$ *in* $\mathcal{M}$ for those directions for which $\lambda^A\xi^{A'}$ is *tangent* to $\mathcal{S}$ by virtue of the 2-surface twistor equation on $\mathcal{S}$. Third, note that two neighbouring points of $\mathcal{S}$ are null separated if and only if their images in $\mathbf{CM}^{\#}(\mathcal{S})$ are null separated; for if neighbouring points $p, p' \in \mathcal{S}$ satisfy $(pp')^{AA'}\lambda_A = 0$, then (p, λ_A) and (p', λ_A) both annihilate the same $\{\omega^A\}$, namely one given by $\omega^A\lambda_A = 0$ at p and therefore at p' since $\lambda^A\xi^{A'}\nabla_{AA'}(\omega^B\lambda_B) = 0$. Thus, the *complex conformal metric of* $\mathcal{S}$ *is preserved* by its embedding in $\mathbf{CM}^{\#}(\mathcal{S})$. Now the spin-vector fields ω^A in $\mathbf{CM}^{\#}(\mathcal{S})$ for which $\omega^A\lambda_A$ is constant along the $\lambda^A\xi^{A'}$ directions in

$\mathbf{CM}^{\#}(\mathcal{S})$ induce spin-vector fields on $\mathcal{S}$ which are constant along those same null directions at $\mathcal{S}$ when they happen to be tangent to $\mathcal{S}$ in $\mathcal{M}$, which, by the above, gives the 2-surface twistor equation.

In $\mathbf{CM}^{\#}(\mathcal{S})$, the 2-surface twistor equation takes the standard form

$$\eth'\omega^0 = \sigma'\omega^1, \qquad \eth\omega^1 = \sigma\omega^0 \tag{4}$$

with respect to a spinor basis associated with the tangent 2-plane element, and with respect to some complex flat metric for $\mathbf{CM}(\mathcal{S})$. These σ, σ' must scale in the standard way under conformal, and dyad, rescalings—which amounts to:

$$\sigma \mapsto \Gamma^2 \tilde{\Gamma}' \Gamma'^{-1} \sigma, \qquad \sigma' \mapsto \Gamma'^2 \tilde{\Gamma} \Gamma^{-1} \sigma' \tag{5}$$

under

$$o^A \mapsto \Gamma o^A, \qquad \iota^A \mapsto \Gamma' \iota^A, \qquad \tilde{o}^{A'} \mapsto \tilde{\Gamma} \tilde{o}^{A'}, \qquad \tilde{\iota}^{A'} \mapsto \tilde{\Gamma}' \tilde{\iota}^{A'}$$

so with normalized dyads,

$$\varepsilon_{AB} \mapsto (\Gamma\Gamma')^{-1} \varepsilon_{AB}, \qquad \varepsilon_{A'B'} \mapsto (\tilde{\Gamma}\tilde{\Gamma}')^{-1} \varepsilon_{A'B'}.$$

Comparing (4) in $\mathbf{CM}^{\#}(\mathcal{S})$ with (4) in $\mathcal{M}$, we find that the corresponding σ, σ' must scale, as in (5), between the two as required. More can be done using these ideas—later!

156
A SUGGESTED FURTHER MODIFICATION TO THE QUASI-LOCAL FORMULA

Twistor Newsletter **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

A suggested further modification to the quasi-local formula. *Twistor Newsletter* **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.20 **A suggested further modification to the quasi-local formula** *by R.Penrose* (**TN** 20, September 1985)

In view of the anomalous expression for the quasi-local mass for small contorted ellipsoids found by N.M.J.Woodhouse (§II.4.16), it seems that one must seek a further modification beyond that given by the determinant factor as suggested in §II.4.15. An appealing feature of the *original* form had always been the existence of the Tod expression $-\frac{i}{2\pi G}\oint \pi_{0'}\pi_{1'}\underline{\mathcal{Y}}$ ($\underline{\mathcal{Y}}$ being the surface-area 2-form). The determinant modification

$$-\tfrac{i}{4\pi G}\oint\{(\Psi_1-\Phi_{01})(\omega^0)^2+(\Psi_2-\Phi_{11}-\Lambda)2\omega^0\omega^1+(\Psi_3-\Phi_{21})(\omega^1)^2\}\eta\underline{\mathcal{Y}}$$

with $\eta\propto\begin{vmatrix} \omega^0_1 & \omega^0_2 & \omega^0_3 & \omega^0_4 \\ \omega^1_1 & \cdots & \cdots & \cdots \\ \partial'\omega^0_1 & \cdots & \cdots & \cdots \\ \partial'\omega^1_1 & \cdots & \cdots & \cdots \end{vmatrix}$ leads, by integration by parts, to a 'Tod form'

$$-\tfrac{i}{2\pi G}\oint(\hat{\pi}_{0'}\hat{\pi}_{1'}+\omega^0\omega^1\partial\nu\partial'\nu)\underline{\mathcal{Y}}$$

where $\hat{\pi}_{0'}=\nu\pi_{0'}+i\omega^1\partial'\nu$, $\hat{\pi}_{1'}=\nu\pi_{1'}+i\omega^0\partial\nu$ and $\nu^2=\eta$. These modifications to $\pi_{0'}$, $\pi_{1'}$ seem natural enough as 'complex conformal factor' terms, but the extra term $\omega^0\omega^1\partial\nu\partial'\nu$ has an awkward appearance. Though the motivation is a little flimsy, it would seem to be worth trying

$$-\tfrac{i}{2\pi G}\oint\hat{\pi}_{0'}\hat{\pi}_{1'}\underline{\mathcal{Y}}$$

as a possible modified definition for $A_{\alpha\beta}Z^\alpha Z^\beta$. The *good* results to date seem to be unaffected by this change. Thanks to K.P.Tod, N.M.J.Woodhouse for comments.

Reference

Penrose, R. & Rindler, W. (1986) Spinors and Space-Time Vol. 2, C.U.P.

157
HIGHER-DIMENSIONAL TWO-SURFACE TWISTORS

Twistor Newsletter **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

 Higher-dimensional two-surface twistors. *Twistor Newsletter* **20** (1985). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.21 **Higher-dimensional two-surface twistors** *by R.Penrose* (**TN** 20, September 1985)

Generalizations of the 2-surface twistor concept to higher dimensions can be achieved in two different directions. In the first place we can keep the surface $\mathcal{S}$ two-dimensional, but allow the ambient space to become higher-dimensional. In the second place we can allow $\mathcal{S}$ itself to become higher-dimensional. In any case, the spinors and twistors need to be those for a higher-dimensional ambient space-time $\mathcal{M}$. For simplicity, I shall consider $\mathcal{M}$ to be 6-dimensional and $\mathcal{S}$ to be either 2-dimensional (Case A) or 4-dimensional (Case B), but higher generalizations can also be readily achieved.

Here the reduced spinors for $\mathcal{M}$ are 4-dimensional and the twistors for $\mathcal{M}$, if $\mathcal{M}$ happens to be conformally flat, would be 8-dimensional (i.e. reduced spinors for the pseudo-orthogonal group for which the points of $\mathcal{M}$ correspond to generators of the null cone; see the Appendix to Penrose & Rindler (1986) for the relevant n-dimensional spinor and twistor concepts). For the moment, at least, everything is complex, so the signature is irrelevant.

We need to see how to describe, in spinor terms, the projection operator E which takes us down from the 6-space to the 2-space (A) tangential to $\mathcal{S}$ or (B) normal to $\mathcal{S}$ and its orthogonal complement $F = I - E$ which takes us from the 6-space to the 4-space normal to $\mathcal{S}$ (case A)

tangential to $\mathcal{S}$ (case B).

We use a 'twistor-type' description of vectors in $\mathcal{M}$, where lower case Greek indices are 4-dimensional ('twistor-like') spinor indices, skew pairs of which represent tangent vectors in $\mathcal{M}$. The tangent space (Case A) or normal space (Case B) to $\mathcal{S}$ at a point p ($\mathcal{S}$ assumed non-null) contains two (complex) null directions U, V. Now null vectors in $\mathcal{M}$ correspond to simple bivectors in the (twistor-like) spin space of $\mathcal{M}$:

$$U^{\alpha\beta} = a^{[\alpha}b^{\beta]}, \qquad V^{\alpha\beta} = c^{[\alpha}d^{\beta]}$$
$$U_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}a^{\gamma}b^{\delta}, \qquad V_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}c^{\gamma}d^{\delta}$$

and take $a^{\alpha}b^{\beta}c^{\gamma}d^{\delta}\varepsilon_{\alpha\beta\gamma\delta} = 4$. The projector E is then given by

$$E^{\alpha\beta}_{\gamma\delta} = \tfrac{1}{2}U^{\alpha\beta}V_{\gamma\delta} + \tfrac{1}{2}V^{\alpha\beta}U_{\gamma\delta}$$

and

$$F^{\alpha\beta}_{\gamma\delta} = \delta^{[\alpha}_{\gamma}\delta^{\beta]}_{\delta} - E^{\alpha\beta}_{\gamma\delta}.$$

It is convenient to consider corresponding projectors in the spin-space:

$$e^{\alpha}_{\beta} = U^{\alpha\beta}V_{\beta\gamma}, \qquad f^{\alpha}_{\beta} = V^{\alpha\gamma}U_{\beta\gamma}$$

which decompose the spin-space:

$$\delta^{\alpha}_{\beta} = e^{\alpha}_{\beta} + f^{\alpha}_{\beta} \tag{A}$$

where

$$e^{\alpha}_{\beta}e^{\beta}_{\gamma} = e^{\alpha}_{\gamma}, \quad f^{\alpha}_{\beta}f^{\beta}_{\gamma} = f^{\alpha}_{\gamma}, \quad e^{\alpha}_{\beta}f^{\beta}_{\gamma} = 0 = f^{\alpha}_{\beta}e^{\beta}_{\gamma}.$$

Note also

$$e^{[\alpha}_{\gamma}e^{\beta]}_{\delta} = \tfrac{1}{2}U^{\alpha\beta}V_{\gamma\delta}, \quad f^{[\alpha}_{\gamma}f^{\beta]}_{\delta} = \tfrac{1}{2}V^{\alpha\beta}U_{\gamma\delta}. \tag{B}$$

The corresponding decomposition of the vector space is given by

$$\begin{array}{lcccc} \boxed{\text{Case A}}\!:\!\rightarrow \text{normal} & & & \text{tangential} & \\ \delta^{[\alpha}_{\gamma}\delta^{\beta]}_{\delta} & = & \overbrace{2e^{[\alpha}_{[\gamma}f^{\beta]}_{\delta]}} & + & \overbrace{e^{[\alpha}_{\gamma}e^{\beta]}_{\delta} + f^{[\alpha}_{\gamma}f^{\beta]}_{\delta}} \\ \boxed{\text{Case B}}\!:\!\rightarrow \text{tangential} & & & & \text{normal} \end{array} \tag{C}$$

Case A. Let $\mathcal{S}$ be 'spacelike' with S^2 topology. The twistor equation in (conformally flat) 6-space can be written

$$\nabla_{\alpha\beta}\omega^\gamma = \delta^\gamma_{[\alpha}\pi_{\beta]} \tag{D}$$

By (C), $e^{[\cdot}_{\cdot}e^{\cdot]}_{\cdot}$ and $f^{[\cdot}_{\cdot}f^{\cdot]}_{\cdot}$ project tangentially to $\mathcal{S}$ and we can extract the tangential parts of (D) as

$$e^\lambda_\gamma f^\alpha_\mu f^\beta_\nu \nabla_{\alpha\beta}\omega^\gamma = 0, \qquad f^\lambda_\gamma e^\alpha_\mu e^\beta_\nu \nabla_{\alpha\beta}\omega^\gamma = 0$$

i.e. (by B)

$$e^\lambda_\gamma V^{\alpha\beta}\nabla_{\alpha\beta}\omega^\gamma = 0, \qquad f^\lambda_\gamma U^{\alpha\beta}\nabla_{\alpha\beta}\omega^\gamma = 0$$

i.e.

$$\underbrace{e^\mu_\lambda V^{\alpha\beta}\nabla_{\alpha\beta}(\omega^\gamma e^\lambda_\gamma)}_{\partial' \quad \omega^0} = \underbrace{\omega^\gamma f^\lambda_\gamma\, e^\mu_\nu V^{\alpha\beta}\nabla_{\alpha\beta}e^\nu_\lambda}_{\omega^1 \quad \sigma'} \tag{E}$$

analogue:

$$\underbrace{f^\mu_\lambda U^{\alpha\beta}\nabla_{\alpha\beta}(\omega^\gamma f^\lambda_\gamma)}_{\partial \quad \omega^1} = \underbrace{\omega^\gamma e^\lambda_\gamma\, f^\mu_\nu U^{\alpha\beta}\nabla_{\alpha\beta}f^\nu_\gamma}_{\omega^0 \quad \sigma}$$

analogue:

Similar use of the Atiayh-Singer theorem to that used for the original S^2 in 4-space applies here also. (Look at a canonical case, with its adjoint, and then deform.) In general we get an 8-dimensional space of solutions of (E).

Case B. Let $\mathcal{S}$ become a 4-dimensional space-time $\mathcal{M}$ of embedding class 2 (or 1), $\mathcal{M}$ being now a conformally flat 6-space. For example $\mathcal{S}$ could be the Schwarzschild solution $\mathcal{M}_S$. The restriction of (D) to $\mathcal{M}$ turns out to take the form

$$\tilde\nabla^{A'}_{(A}\omega_{B)} = \Gamma^{A'B'}_{AB}\tilde\omega_{B'}, \qquad \tilde\nabla'^{A}_{(A'}\tilde\omega_{B')} = \tilde\Gamma^{AB}_{A'B'}\omega_B$$

where $\tilde\nabla_{AA'}$ and $\tilde\nabla'_{AA'}$ are analogues of ∂ and ∂', which differ from $\nabla_{AA'}$ by 'gauge' dependent terms - like α, β, γ, ϵ of the spin-coefficient-formalism. Strictly ω_A and $\tilde\omega_{A'}$ are 'boost-weighted' spinors on $\mathcal{M}$, the boost weight referring to the perpendicular frame transformations. Thanks to Toby Bailey, Lane Hughston, Paul Tod and Ian Gatenby.

References

Penrose, R. & Rindler, W. (1986) Spinors and space-time. vol. 2, C.U.P.

158
NUMERICAL RELATIVITY BY POWER SERIES

Galaxies, axisymmetric systems and relativity. Essays presented to W. B. Bonnor on his 65th Birthday. Ed. M. A. H. MacCallum. Cambridge University Press, 1985

This is a proposal for a very non-standard approach to numerical general relativity, which is free of the standard coordinate-freedom problems. No doubt it would have many other sorts of problem of its own.

404 Numerical relativity by power series. *Galaxies, axisymmetric systems and relativity. Essays presented to W. B. Bonnor on his 65th Birthday.* Ed. M. A. H. MacCallum. 1985

NUMERICAL RELATIVITY BY POWER SERIES

R. Penrose
Mathematical Institute, Oxford

Exact solutions of Einstein's equations may sometimes be be found by good fortune, or else, as has been Bill Bonnor's way (e.g. Bonnor 1966, 1977), by rare combinations of ingenuity and power of technique (see also Kramer, Stephani, MacCallum and Herlt 1980). But for a solution for some given physical problem to have a reasonable closed-form expression at all, one is very much at the mercy of fortune.

Accordingly, many relativists have turned their attention to the methods of numerical relativity (cf. Smarr 1979). Here one seems to have the freedom, for any given physical problem, to specify initial data on some spacelike hypersurface $\mathcal{H}$ and the evolution according to Einstein's equations is carried out numerically. However, as is well appreciated, there are a great many difficulties in this procedure. In the first place, there is the considerable complexity of the equations themselves, in the full four dimensions, so that symmetry or other restrictions have normally to be imposed in order that the equations can be simplified sufficiently to be handled at all adequately by even the most powerful of modern computers. There is the problem, also, that the data on $\mathcal{H}$ is subject to differential constraints and cannot simply be chosen freely. One must first solve the (elliptic but non-linear) constraint equations on $\mathcal{H}$, say by some method which "improves" an initial guess for the appropriate data at $\mathcal{H}$. Then, as the system is evolved this "improvement" process must be re-applied from time to time because although the exact field equations preserve satisfaction of the constraint equations, the approximations involved in the numerical evolution will continually introduce new errors. Coordinate conditions must, of course, be imposed in order that the evolution be determinate. But one must also be careful to choose these conditions so that the resulting propagation of the coordinates is not in some sense unstable (e.g. the coordinate surfaces might "tilt over" and become null, narrow down and fail to cover some region of the space-time, expand rapidly outwards to infinity, or simply produce spurious coordinate singularities). Furthermore, when a metric has indeed been successfully computed, there is the problem of interpreting what it means. The curvature tensor must be computed, and for this it is rather more illuminating to have the components in some (say orthonormal or null) tetrad rather than in coordinate form. But then, some automated means of specifying the tetrad at each point must also be given.

All these (and other) problems make it clear that numerical computation in general relativity, at least in its standard form, is a formidable undertaking. My purpose here is to suggest an alternative procedure of a completely different nature. I have no doubt that my own proposal will be found to have formidable problems of its own - and perhaps one or more of these will render it impracticable - but at least none of the problems that I have mentioned as arising in the standard procedure are particularly relevant with it.

What I have in mind is related to an approach based on the characteristic initial value problem for general relativity (compare Corkill and Stewart 1983) - i.e. where $\mathscr{H}$ is taken to become a null hypersurface $\mathscr{N}$. Such an approach has the advantage that the constraints on the initial data effectively disappear, an irredundant locally defined and freely specifiable quantity (see Penrose 1980, Penrose and Rindler 1984), referred to as the null-datum Ψ (or equivalently the complex shear σ or intrinsic degenerate conformal metric), completely determining the space-time $\mathscr{M}$ (in the domain of dependence) in the vacuum case. In 2-spinor terms we have

$$\Psi = \Psi_{ABCD}\, \xi^A \xi^B \xi^C \xi^D \qquad (1)$$

where Ψ_{ABCD} is the (symmetric) Weyl conformal spinor and ξ^A is a spin-vector whose flagpole points along the generators of $\mathscr{N}$ (and which, for convenience, may be chosen to be parallelly propagated along these generators). (See Penrose and Rindler 1984 for a comprehensive treatment of the 2-spinor formalism.)

A problem with the use of a null initial data hypersurface has frequently been pointed out. This is that caustics and crossing regions arise and one has to worry about consistency conditions with the data at such regions in addition to severe problems concerning the propagation there. However, in the approach that I am suggesting, it seems that caustics and crossing regions do not in themselves cause any difficulties.

My proposal is based on a procedure I suggested many years ago (Penrose 1960), and it depends upon showing that, in vacuum, the various covariant derivatives

$$\Psi_{ABCD},\ \nabla^{E'}_{E}\Psi_{ABCD},\ \nabla^{F'}_{F}\nabla^{E'}_{E}\Psi_{ABCD},\cdots \qquad (2)$$

are completely determined by the symmetrized derivatives

$$\Psi_{ABCD},\ \nabla^{E'}_{(E}\Psi_{ABCD)},\ \nabla^{(F'}_{(F}\nabla^{E')}_{E}\Psi_{ABCD)},\ \cdots \qquad (3)$$

the latter being freely specifiable (apart from certain "growth" restrictions) at any given point $P \in \mathscr{M}$. The quantities (3) at P provide the coefficients for a power series development of the null-datum Ψ on $\mathscr{N}$, where $\mathscr{N}$ is now specified as being the light cone of P. The quantities (2) at P provide the coefficients for a power series

development of the full vacuum curvature Ψ_{ABCD} of $\mathcal{M}$.

My suggestion is that a "once and for all" algebraic computation is first performed for finding the explicit expressions of (2) in terms of (3), up to as many derivatives as can reasonably be achieved on some very powerful computer. This would be by far the most difficult part of the procedure. I have no idea how many derivatives could be handled in this way: perhaps not very many, since the explicit procedures seem to involve a rapidly expanding number of operations. No doubt things can be kept somewhat under control with the exercise of ingenuity. I imagine that if as many as about ten derivatives of Ψ_{ABCD} could be worked out for the general case then the whole procedure would be well worth while.

The (approximate) calculation of a specific space-time would, given the above, be a comparatively simple matter. The input data would be the quantities (3), at P, chosen to provide the power series development of some preassigned geometry for a light cone. The components would always be taken with respect to some fixed spin-frame (o^A, ι^A) at P. For n derivatives we should have $5 + 6 \times 2 + 7 \times 3 + \ldots + (n+5)(n+1) = \frac{1}{6}(n+1)(n+2)(2n+15)$ complex components. Then the quantities (2) would be worked out and used to calculate the power series for the curvature Ψ_{ABCD} at a general point Q. The point Q would be located with respect to P by specifying a vector v^a at P to provide the tangent direction to a geodesic PQ, and whose length (extent) would be equal to that of PQ. The computation provides the components of Ψ_{ABCD} at Q with respect to a spin-frame which is the given one at P parallelly propagated along PQ.

Note that the curvature is calculated directly in this method, and in a form where the invariants and Petrov type may be immediately obtained (Penrose 1960, Penrose and Rindler 1984). Moreover, it seems not to matter if caustics or crossing regions are encountered in the exponential map. This is exemplified by the case of a plane wave, for which the entire space-time may be explicitly obtained by this method without any difficulty (Penrose 1960) despite the fact that caustics and crossing regions do occur in the exponential map.

It would be easy to set up a situation in which trapped surfaces can be found on $\mathcal{N}$. For example, we could take

$$\Psi_{ABCD} = o_A o_B o_C o_D, \quad \nabla^{E'}_{(E} \Psi_{ABCD)} = \iota^{E'} \iota_E \iota_A \iota_B \iota_C \iota_D \qquad (4)$$

at P, with the remaining quantities in (3) zero. It would be interesting to get some understanding of how the curvature behaves near the resulting singularity. Of course divergences can appear with a power series even when the quantity in question is not itself actually blowing up. However, there are methods for distinguishing a genuine divergence from "spurious" ones that can arise in this way. But one should also bear in mind that only a finite number of terms in the series are available in any case, so some estimates are needed so that one can judge the

trustworthiness of the results obtained. It may well be that simply summing the series is not the most effective way of obtaining the best estimate of the curvature along PQ and some other procedure may be preferable when divergences are encountered (e.g. the use of Padé approximates, etc.). There is also the possibility, when "spurious" divergences arise, that one could proceed in steps: from P to R, then from R to Q, etc., in the manner of analytic continuation.

It seems to me that whatever the eventual merits, or demerits, of this prcoedure may be, it would at least be very interesting to experiment with it and see how far it will go.

The author is grateful to members of the Relativity Group at the University of Waterloo for enlightening discussions.

References

Bonnor, W.B. (1966) Z. Phys. 190, 444.

Bonnor, W.B. (1977) J. Phys. A 10, 1673.

Corkill, R.W. and Stewart, J.M. (1983) Proc. R. Soc. Lond. A386, 373.

Kramer, D., Stephani, H., MacCallum, M. and Herlt, E. (1980) Exact Solutions of Einstein's Field Equations (VEB Deutscher Verlag der Wissenschaften, Berlin and Cambridge Univ. Press, Cambridge).

Penrose, R. (1960) Ann. Phys. 10, 171 .

Penrose, R. (1980) Gen. Rel. Grav. 12, 225.

Penrose, R. and Rindler, W. (1984) Spinors and Space-Time; Vol. 1: Two-Spinor Calculus and Relativistic Fields (Cambridge Univ. Press, Cambridge).

Smarr, L. (1979) in Sources of Gravitational Radiation (ed. L. Smarr; Cambridge Univ. Press, Cambridge).

159
ELEMENTS OF TWISTOR THEORY

(Based on a Lecture Delivered in Chittagong 1986, Bangladesh). *Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong.* Ed. J. N. Islam, 1986

410 Elements of twistor theory. (Based on a Lecture Delivered in Chittagong 1986, Bangladesh). *Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong.* Ed. J. N. Islam, 1986

Elements of Twistor Theory*

Roger Penrose

Mathematical Institute, Oxford, United Kingdom
and
Institute for Theoretical Physics, University of California,
Santa Barbara, California 93106, U.S.A.

* Lecture given at International Conference on Theoretical Physics, University of Chittagong, Chittagong, Bangladesh, in January 1986.

Elements of twistor theory. (Based on a Lecture Delivered in Chittagong 1986, Bangladesh). 411
Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong. Ed. J. N. Islam, 1986

Twistor theory provides a mathematical way of looking at space-time structure which is quite different from the standard way. Instead of space-time points (or events) being the primary objects out of which the space is to be constructed, these points are regarded as secondary objects, built up from things, referred to as *twistors*, which are regarded as more primitive. The idea is that all physical concepts are to be defined directly in terms of twistors, and this often leads to some remarkable mathematical expressions. The status of the theory at present is that, on the one hand, ther are many surprising expressions of this kind, coding the information of the solutions of physical equations (frequently nonlinear) but, on the other hand, much more needs to be done to find twistor transcriptions of other equations which have so far stubbornly refused to yield completely, despite various encouraging signs.

As a first step toward understanding the physical or geometrical meaning of the twistor concept, imagine a light ray, (*i.e.*, null straight line) in Minkowski space $\mathbf{M}$. As we shall see shortly, such a light ray can be interpreted as what will be referred to as a *projective null twistor*. The space of such light rays (including a 2-parameter family of idealized rays lying entirely at infinity) will be denoted $\mathbf{PN}$. The idea will be to regard the space $\mathbf{PN}$ as more primitive than $\mathbf{M}$. A space-time point $X \in \mathbf{M}$ is thus interpreted as a locus in $\mathbf{PN}$, namely that locus which corresponds to the family of light rays through X, *i.e.*, the generators of the light cone of X. This locus is a sphere S^2, and it is significant to regard this sphere as a *Riemann sphere* — *i.e.*, the Argand plane of complex numbers, together with a point representing ∞. The holomorphic (complex-analytic) motions of this sphere to itself correspond precisely to the (restricted) *Lorentz group* acting at the point X. Generalizing this idea, we are led to hope that the entire space $\mathbf{PN}$, not merely these Riemann spheres, might have an interpretation in terms of holomorphic structure. However, $\mathbf{PN}$ itself cannot be a complex manifold, being 5-dimensional as a real space. On the other hand, it can be completed to a 6-real-dimensional space $\mathbf{PT}(\supset \mathbf{PN})$ which

does indeed have a natural complex structure, namely as a complex projective 3-space. We shall see in a moment that the points of **PT** not lying on **PN** have interpretations as (classical) free photons with intrinsic spin. The extra dimension comes about because now the energy of the photon has to be taken into consideration.

To make all this more precise, we shall need to use a 2-spinor formulation. Recall that the 2-spinor description $V^{AA'}$ of a 4-vector V^a is, in terms of components with respect to an orthonormal frame and associated spin-frame,

$$\begin{bmatrix} V^{00'} & V^{01'} \\ V^{10'} & V^{11'} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} V^0 + V^3 & V^1 + iV^2 \\ V^1 - iV^2 & V^0 - V^3 \end{bmatrix},$$

the vector being null whenever this matrix has rank one. (The space-time signature is $+---$ and units are such that $c = 1$.) Now consider a massless particle with 4-momentum p_a and angular momentum $M^{ab} (= -M^{ba})$ with respect to some origin O. The requirements for a *massless* particle are, first, that p_a be future-null and, second, that the Pauli-Lubanski spin 4-vector

$$S_a = \frac{1}{2} \epsilon_{abcd} p^b M^{cd}$$

(where ϵ_{abcd} is the Levi-Civita tensor) be proportional to the 4-momentum:

$$S_a = s p_a$$

These conditions can be encoded in spinor form as

$$p_{AA'} = \bar{\pi}_A \pi_{A'}$$

and

$$M^{AA'BB'} = i\omega^{(A} \bar{\pi}^{B)} \epsilon^{A'B'} - i\epsilon^{AB} \bar{\omega}^{(A'} \pi^{B')}.$$

Here round brackets denote symmetrization, the ϵ symbols denote Levi-Civita spinors (used for raising and lowering spinor indices), and the bar denotes complex conjugation (this operation interchanging primed with unprimed indices). (For further details see Penrose and Rindler 1984, 1986, Huggett and Tod 1985.)

The two spinors of the pair $(\omega^A, \pi_{A'})$ are the two *spinor parts* of the *twistor* Z^α, and we sometimes write

$$\omega^A = Z^A, \quad \pi_{A'} = Z_{A'},$$

the four-dimensional upper twistor index α being regarded as a direct sum of an upper unprimed spinor index and a lower primed one. The twistor Z^α is uniquely determined, up to a phase transformation

$$Z^\alpha \mapsto e^{i\theta} Z^\alpha, \quad i.e. \quad (\omega^A, \pi_{A'}), \mapsto (e^{i\theta}\omega^A, e^{i\theta}\pi_{A'}), \ (\theta \in \mathbf{R})$$

by the momentum and angular momentum of the massless particle. Conversely, any twistor for which $\pi_{A'} \neq 0$ (up to such a phase freedom) uniquely determines such a momentum and angular momentum. The complex conjugate of the twistor Z^α is the *dual* twistor $\bar{Z}_\alpha$, whose spinor parts constitute the pair

$$(\bar{\pi}_A, \bar{\omega}^{A'})$$

the twistor "norm" $Z^\alpha \bar{Z}_\alpha$ turning out to be (twice) the *helicity* s that arose above as the proportionality factor betwen the spin 4-vector and the momentum 4-vector:

$$Z^\alpha \bar{Z}_\alpha = \omega^A \bar{\pi}_A + \pi_{A'} \bar{\omega}^{A'} = 2s.$$

Since M^{ab} depends on a choice of origin, so does ω^A. We may move to an arbitrary point $X \in \mathbf{M}$ and take moments about *that* point. We then find, after some calculation, that

$$\omega^A(X) = \omega^A(O) - ix^{AA'}\pi_{A'}(O)$$

where we choose the phase so that $\pi_{A'}(X)$ is constant:

$$\pi_{A'}(X) = \pi_{A'}(O).$$

(Here $x^{AA'}$ is the position vector of X with respect to O, and the bracketed symbols refer to the choice of origin.) Regarding $\omega^A(X)$ as a kind of "field," we find that the above transformation properties are precisely encoded in the equation

$$\nabla^{(A}_{A'}\omega^{B)}(X) = 0,$$

which is referred to as the *twistor equation*, where the constant $\pi_{A'}$ is determined by the ω^A-field according to

$$\nabla_{AA'}\omega^B(X) = -i\epsilon_A{}^B\pi_{A'}.$$

The gradient operator $\nabla_{AA'}$, with respect to a constant coordinate frame, is just $\partial/\partial x^{AA'}$.

When the helicity vanishes, we refer to Z^α as a *null twistor*. The subset of $\mathbf{T}$ consisting of null twistors is denoted by $\mathbf{N}$. It turns out that for such twistors, the locus of points X for which $\omega^A(X)$ vanishes is a light ray z. (This assumes $\pi_{A'} \neq 0$. When $\pi_{A'} = 0$ we get an idealized light ray at infinity.) Conversely, any light ray in $\mathbf{M}$ determines a null twistor up to the freedom

$$Z^\alpha \mapsto \lambda Z^\alpha \quad (\lambda \in \mathbf{C}, \lambda \neq 0).$$

The proportionality classes of twistors up to such a scaling freedom are referred to as *projective twistors*, and the space of all such projective twistors is denoted by $\mathbf{PT}$. Correspondingly, $\mathbf{PN}$ denotes the space of projective null twistors, as indicated earlier. Note that the light ray z is the set of points about which the angular momentum vanishes, so we may say that z is the world-line of the massless, spinless particle described by a null twistor.

When the twistor is not null, there are no points of $\mathbf{M}$ about which the angular momentum vanishes. Moreover, there is no real world-line that is singled out in a Poincaré invariant way by the momentum and angular-momentum structure. (This is in stark contrast with the case of a massive particle.) Accordingly, it is often convenient to pass

to a description in terms of the complexification CM of M, where the position vector x^a is allowed to become a complex vector. This means that the spinor description $x^{AB'}$ need not be Hermitian, but the same formula relating $\omega^A(X)$ to $\omega^A(O)$, $x^{AA'}$ and $\pi_{A'}$ holds, as before. The locus within CM at which $\omega^A(X)$ vanishes is referred to as an *α-plane* (including the limiting cases where $\pi_{A'} = 0$, provided that $Z^\alpha \neq 0$). The α-planes are two-complex-dimensional loci on which the induced metric vanishes completely, *i.e.*, the complex distance between any two points on an α-plane is zero. Such planes are sometimes called *totally null*. There are two families of totally null planes in CM, the α-planes and the β-planes. The α-planes have tangent vectors of the form $\lambda^A\pi^{A'}$ where $\pi^{A'}$ is fixed and λ^A varies; the β-planes have tangent vectors of this same form but where λ^A is fixed and it is $\pi^{A'}$ that varies. A dual twistor W_α (up to proportionality) determines a β-plane. The β-planes are complex conjugates of α-planes, where $\bar{Z}_\alpha$ determines that β-plane which is conjugate to the α-plane determined by Z^α.

In the projective space PT, there are three kinds of linear projective subspaces, namely the points, lines and planes. It follows from the previous descriptions that these are represented, in CM, by α-planes, points and β-planes, respectively. We say that one space is *indicent* with another when the dimension of their intersection is greater than in the generic case. Incidence betwen points, lines and planes in PT translates to incidence between the corresponding α-planes, points and β-planes in CM — except that the *meeting* of two lines in PT translates *null-separation* between the corresponding points CM. Two α-planes or two β-planes will generically meet in a single point; an α-plane and a β-plane generically will not intersect at all, but when they do, they intersect in a (complex) *null straight line* in CM (possibly at infinity), and this corresponds, in PT, to a point lying on a plane.

The programme of twistor theory consists of translating, or attempting to translate physical concepts into entirely twistor-space terms. In particular, since it is intended

that the complex numbers which describe twistor space should be closely interrelated with the complex probability amplitudes of quantum theory, one anticipates that quantum wave functions should have a convenient complex description in terms of twistor-space variables. Thus, instead of using the position-space variables x^a — or else the momentum-space variables p_a — which are related to each other via the commutation rules

$$[x^a, x^b] = 0,\ [x^a, p_b] = -i\hbar\delta^a_b,\ [p_a, p_b] = 0,$$

we use Z^α — or else $\bar{Z}_\alpha$ — related by commutation rules

$$[Z^\alpha, Z^\beta] = 0,\ [Z^\alpha, \bar{Z}_\beta] = \hbar\delta^\alpha_\beta,\ [\bar{Z}_\alpha, \bar{Z}_\beta] = 0.$$

One motivation for this choice of twistor commutation rules is that the correct Poincaré group commutators between p_a and M^{ab} are then a consequence, where p_a and M^{ab} are constructed from ω^A and $\pi_{A'}$ in the explicit way described above. If we use the Z^α-description, then we have a twistor wave function $f(Z^\alpha)$ "independent" of $\bar{Z}_\alpha$, where this independence is expressed as

$$\frac{\partial f}{\partial \bar{Z}_\alpha} = 0,$$

i.e., f is *holomorphic* in Z^α. If, alternatively, we were to use the $\bar{Z}_\alpha$-description, we should have a twistor wave function $\bar{f}(\bar{Z}_\alpha)$ which is *anti-holomorphic* in $\bar{Z}_\alpha$, *i.e.*, holomorphic in a *dual* twistor W_α (where $W_\alpha = \bar{Z}_\alpha$).

The *helicity* operator s must now be given by

$$s = \frac{1}{4}(Z^\alpha \bar{Z}_\alpha + \bar{Z}_\alpha Z^\alpha).$$

In terms of the Z^α-description, we must represent

$$\bar{Z}_\alpha \quad \text{by} \quad -\hbar\frac{\partial}{\partial \bar{Z}_\alpha}$$

and in terms of the $\bar{Z}_\alpha$-description we represent

$$Z^\alpha \quad \text{by} \quad \hbar\frac{\partial}{\partial \bar{Z}_\alpha}.$$

Thus, in the first case we have, for the helicity operator,

$$-\frac{1}{2}\hbar\left(Z^\alpha\frac{\partial}{\partial Z^\alpha}+2\right)$$

and in the second case, for this same operator (and writing W_α for $\bar{Z}_\alpha$) we have

$$\frac{1}{2}\hbar\left(W_\alpha\frac{\partial}{\partial W_\alpha}+2\right).$$

We see that the wave functions which are eigenstates of helicity must therefore be *homogeneous* holomorphic functions of a twistor (or dual twistor) — because $Z^\alpha\partial/\partial Z^\alpha$ (or $W_\alpha\partial/\partial W_\alpha$) is Euler's homogeneity operator. The most important cases are listed in the table. (Here "graviton" refers to a massless linear spin 2 field.)

	helicity in units of $\hbar$	Z^α-homogeneity	W_α-homogeneity
right-handed graviton	2	-6	2
right-handed photon	1	-4	0
antineutrino	$\frac{1}{2}$	-3	-1
massless scalar	0	-2	-2
neutrino	$-\frac{1}{2}$	-1	-3
left-handed photon	-1	0	-4
left-handed graviton	-2	2	-6

Notice that whether we adopt the Z^α-description or the W_α-description we have an *asymmetrical* behavior with respect to interchange of left and right. (It would not make sense to use the Z^α-description for some helicities and the W_α-description for others. That would be like using a position description for some kinds of particle while using a momentum description for others.) Thus, the choice of one or other of the Z^α- or W_α-descriptions involves some sort of preference with regard to left- or right-handedness. It

turns out that at the level of non-interacting massless fields, this choice is not of great consequence. However, when interactions are involved the asymmetry is more serious, and the twistor formalism could well lead to a space-reflection asymmetry in the physical processes that it most naturally describes. In the case of weak interactions this could very possibly be a definite advantage. However, with space-reflection-symmetric interactions — and, most particularly, general relativity — this could present a problem. It still remains to be seen whether full general relativity can be adequately coded into twistor terms.

Let us consider, now, the twistor wave functions for free massless particles. (Massive particles will not be considered here. They can be described in terms of functions of more than one twistor. Internal symmetry groups arise in a natural way in terms of transformations betwen these twistors, but certain problems of interpretations, *etc.* remain. See Hughston 1979, Perjés 1977, 1982 for details.) Our wave function is to be a holomorphic twistor function $f(Z^\alpha)$ (where we arbitrarily choose the Z^α-description) which is homogeneous of some degree, as specified in the table. In fact $F(Z^\alpha)$ must necessarily be singular in certain regions, and we assume that these singularities are arranged appropriately so that the following contour integrals can be non-trivially performed:

$$\oint \pi_{A'} \cdots \pi_{D'} f(Z^\alpha) \pi_{E'} d\pi^{E'} = \phi_{A'\ldots D'}(X),$$

in the positive-helicity case, or

$$\oint \frac{\partial}{\partial \omega^A} \cdots \frac{\partial}{\partial \omega^D} f(Z^\alpha) \pi_{E'} d\pi^{E'} = \phi_{A\ldots D}(X),$$

in the negative-helicity case, or

$$\oint f(Z^\alpha) \pi_{E'} d\pi^{E'} = \phi(X)$$

in the zero helicity case. In each case the contour is to be taken within the space of twistors $Z^\alpha = (\omega^A, \pi_{A'})$ which are incident with the point X, i.e., where

$$\omega^A = ix^{AA'} \pi_{A'}$$

(the spinors parts being taken at the origin O) or, equivalently,

$$\omega^A(X) = 0.$$

We may think of the contour as lying in the Riemann sphere which represents the light rays through X, in the case when X is a real point. However, X does not need to be real; and in general this Riemann sphere would represent the family of α-planes through $X \in \mathbf{CM}$, *i.e.*, the projective line in $\mathbf{PT}$ which represents X. We ask that the singularities of f lie in two disjoint closed sets on this Riemann sphere and we perform our contour integral so as to separate one from the other. We suppose that, as X varies, these closed sets remain disjoint — provided that X lies in the region of $\mathbf{CM}$ in which we require our wave function to be defined in its space-time description. We then find, as may be verified by a simple direct calculation, that the appropriate massless free-field equation

$$\nabla^{AA'}\phi_{A'B'\ldots D'}(X) = 0, \ \nabla^{AA'}\phi_{AB\ldots D}(X) = 0, \quad \text{or} \quad \Box\phi(X) = 0,$$

as the case may be, ***automatically holds*** (where $\Box = \nabla_a\nabla^a$) as X varies, as a result of the contour integral expressions.

When the helicity is $\pm\frac{1}{2}\hbar$ we get the Dirac-Weyl neutrino equation. When it is $\pm\hbar$ we get the Maxwell free-space equations in spinor form, the positive helicity case corresponding to a self-dual (complex) Maxwell tensor

$$^*F_{ab} = iF_{ab}, \quad F_{AA'BB'} = \epsilon_{AB}\phi_{A'B'}$$

and the negative helicity case to an anti-self-dual Maxwell tensor

$$^*F_{ab} = -iF_{ab}, \quad F_{AA'BB'} = \phi_{AB}\epsilon_{A'B'}$$

When the helicity is $\pm 2\hbar$ we get linearlized Einstein sourceless fields, in gauge-invariant form, where the linearlized curvature is, respectively, self-dual

$$^*K_{abcd} = iK_{abcd}, \ K_{AA'BB'CC'DD'} = \epsilon_{AB}\epsilon_{CD}\phi_{A'B'C'D'}$$

or anti-self-dual

$$^*K_{abcd} = -iK_{abcd},\ K_{AA'BB'CC'DD'} = \phi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'}.$$

The duals are defined, in each case, by

$$^*F_{ab} = \frac{1}{2}e_{abcd}F^{cd},\ ^*K_{abcd} = \frac{1}{2}e_{abfg}K^{fg}{}_{cd}.$$

The fact that the massless free-field equations are automatically satisfied in each case, as a consequence merely of the homogeneity condition, is an appealing feature of the twistor descriptions. However, the requirements concerning the singularity regions of f seem to present an awkwardness. In particular, such regions are not Lorentz invariant. Moreover, if one adds two functions for which the singularity regions are appropriate in each case, one may find that the singularities of the *sum* are not appropriate. These difficulties are removed if one adopts a somewhat different viewpoint with regard to twistor functions, a viewpoint which leads us to study them in terms of *sheaf cohomology.* I shall indicate the basic ideas.

Let us, for definiteness, suppose that we are requiring our wave functions, in their space-time versions, to be composed only of *positive frequencies* (which, indeed, is a normal requirement for a wave function). An appropriate way to state this requirement is to demand that the space-time field have a holomorphic extension into the *forward tube*, where the forward tube $\mathbf{CM}^+$ is the set of points $X \in \mathbf{CM}$ whose position vectors have imaginary parts which are past-timelike. It turns out that the points of the forward tube are described, in twistor terms, as the lines lying completely in the region $\mathbf{PT}^+$, which is that part of $\mathbf{PT}$ consisting of projective twistors Z^α for which $Z^\alpha \bar{Z}_\alpha > 0$. To correspond to such a positive-frequency field, we expect, therefore, that our twistor function be defined, in some appropriate sense, on the region $\mathbf{PT}^+$. We find, however, that f is not defined holomorphically over all of $\mathbf{PT}^+$, but has singularities, within the region $\mathbf{PT}^+$, which are

Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong. Ed. J. N. Islam, 1986

contained in two disjoint closed sets A and B. Let us define two open sets

$$U = \mathbf{PT}^+ - A \quad \text{and} \quad V = \mathbf{PT}^+ - B.$$

Then the pair $\{U, V\}$ constitutes and *open covering* of $\mathbf{PT}^+$ since we have

$$U \cup V = \mathbf{PT}^+.$$

The region of definition of the twistor function f is now seen to be the intersection

$$U \cup V.$$

If f could be extended to a function holomorphic on the whole of U then the contour integral would yield a zero field because the contour could always be shrunk to zero on the "U-side" of the Riemann sphere. Similarly, if f could be extended to the whole of V, the result would again be zero. Thus (by linearity), in the general case, if f is replaced according to

$$f \mapsto f + (g - h)$$

where g is holomorphic on U and h is holomorphic on V, the resulting space-time field is unchanged. This provides us with a kind of "gauge freedom" for twistor functions. Rather than thinking of our twistor wave function as just a "function" in the ordinary sense, we think of it as a function holomorphic on $U \cap V$ *modulo* functions which are differences between functions defined on U and functions defined on V.

This idea extends to more general coverings of $\mathbf{PT}^+$, say to a collection of open sets

$$\{U_1, U_2, U_3, \ldots\}$$

where

$$\mathbf{PT}^+ = U_1 \cup U_2 \cup U_3 \cup \ldots$$

and which is *locally finite* — which means that no point of $\mathbf{PT}^+$ is contained in an infinite number of different sets U_i. Instead of having just a single function f, we now have

a holomorphic function f_{ij} assigned to each ordered pair (U_i, U_j) and defined on their intersection

$$U_i \cap U_j.$$

(If this intersection is empty, the definition of f_{ij} is vacuous.) We require the skew-symmetry property

$$f_{ij} = -f_{ji}$$

and, furthermore, that on each triple overlap $U_i \cap U_j \cap U_k$

$$f_{ik} = f_{ij} + f_{jk}$$

Such a collection $\{f_{ij}\}$ is referred to as a (Čech) *cocycle*, and any cocycle which has the form

$$f_{ij} = h_i - h_j,$$

for all i, j, where each h_k is holomorphic over the whole of its corresponding U_k, is referred to as a *coboundary*. Such cocycles *modulo* coboundaries will be referred to as holomorphic 1-functions (or, more conventionally, elements of the first holomorphic sheaf cohomology group) with respect to the covering $\{U_i\}$. We can eliminate the dependence on the covering by showing how such a 1-function may be carried over to finer and finer refinements of the covering and then taking a direct limit. However, in practice it is not necessary to perform such limiting oeprations, there being well-defined procedures for telling when a covering is "fine enough", so that further refinement is not necessary. When adding two 1-functions defined with respect to different coverings, one must take the common refinement of the two coverings. In many situations, however, a specific two-set covering, such as $\{U, V\}$ above, will suffice.

In any case, when we refer to a "twistor function" this will normally be taken to mean a 1-function, as defined by some holomorphic Čech cocycle $\{f_{ij}\}$. Contour integrals can be

directly performed on such families of holomorphic functions. This is done using a concept of "branched contour integral" where the total integral represents a sum of parts, each part being a line integral of an f_{ij} from a point in some $U_i \cap U_j \cap U_k$ to a point in another $U_i \cap U_j \cap U_\ell$. (See Penrose and Rindler (1986) for further details.)

These procedures provide a satisfactory means of describing twistor wave-functions for free massless particles. There is no great difficulty in extending them to give the Fock spaces for many-particle states. Also, some suggestions with regard to massive fields have been made (c.f. Hughston 1979), and there are relations betwen these ideas and the "twistor diagram" approach to expressing interactions (c.f. Hodges and Huggett 1980, Hodges 1982, 1985).

A more specific generalization for handling non-linearities is to the "non-linear 1-functions" that produce actual non-trivial *deformations* of portions of twistor space. Such deformations provide natural ways to express *interactions* in twistor theory. Rather than f_{ij} playing a passive role as a "function" *on* twistor space, it is to be regarded as providing a transition relation patching together two separate pieces of twistor space U_i and U_j. In place of the linear skew-symmetry relation $f_{ij} = -f_{ji}$ we have the fact that the relation expressing the patching of U_i to U_j followed by the patching of U_i to U_j is the inverse of that expressing the patching of U_j to U_i. Likewise, in place of $f_{ik} = f_{ij} + f_{jk}$ on triple overlaps we have the usual functional consistency relation which expresses the fact that the patching of U_j to U_k yields, on the triple overlap, the required patching of U_i to U_k. The *infinitesimal* versions of such consistent patching rules are indeed (linear) 1-functions, in fact, 1-function holomorphic *vector fields*. These vector fields describe the infinitesimal slidings of one patch over another.

As an important example of this, consider a linear 1-function of homogeneity 2, repre-

 Elements of twistor theory. (Based on a Lecture Delivered in Chittagong 1986, Bangladesh). *Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong.* Ed. J. N. Islam, 1986

senting a (linearlized) left-handed gravition. We can construct a vector field on $U_i \cap U_j$:

$$I^{\alpha\beta}\frac{\partial f_{ij}}{\partial Z^\alpha}\frac{\partial}{\partial Z^\beta},$$

where the *infinity twistor* $I^{\alpha\beta}$ is defined by

$$I^{\alpha\beta} = \begin{pmatrix} \epsilon^{AB} & 0 \\ 0 & 0 \end{pmatrix}.$$

Displacement along such a vector fields preserves a certain local structure on the patches U_i, U_j of twistor space (characterized by, in addition to holomorphicity, the differential forms $\pi_{E'}d\pi^{E'}$ and $dZ^0 \wedge dZ^1 \wedge dZ^2 \wedge dZ^3$, and the Euler operator $Z^\alpha\partial/\partial Z^\alpha$) but it does not preserve the linear vector-space structure. If we consider a *finite* deformation $\tilde{\mathbf{T}}$ of (part of) twistor space, possessing this local structure, we can envisage reconstructing a (complex) "space-time" $\widetilde{\mathbf{CM}}$, whose points are appropriate analogues of the projective lines in $\mathbf{PT}$, namely holomorphic curves in $\mathbf{P}\tilde{\mathbf{T}}$ of the correct topological class. Remarkably, it turns out that such holomorphic curves form a (complex) four-dimensional system, so that $\widetilde{\mathbf{CM}}$ is indeed a complex 4-manifold; and, moreover, with a choice of metric determined naturally by the structure of $\tilde{\mathbf{T}}$, the space $\widetilde{\mathbf{CM}}$ automatically satisfies the Einstein vacuum equations for anti-self-dual curvature. In fact the *general* anti-self-dual solution (locally) of the Einstein vacuum equations is obtained in this way. The metric for $\widetilde{\mathbf{CM}}$ arises as a natural generalization of that for $\mathbf{CM}$. In particular, two points of $\widetilde{\mathbf{CM}}$ are null-separated iff the corresponding holomorphic curves in $\mathbf{P}\tilde{\mathbf{T}}$ meet one another, just as was the case for $\mathbf{PT}$. (For further details, see Penrose 1976, Penrose and Ward 1980, Huggett and Tod 1985.)

This shows that our deformation procedure indeed yields the correct non-linear version, (*i.e.*, anti-self-dual general relativity) of the linear negative-helicity massless free field. In order to treat *full* general relativity as a non-linear version of the full spin 2 linear massless thoery, we shall first need an analogous treatment (in the Z^α-description) of the positive-helicity case. This has been referred to as the "googly problem" (because of the cricketing

analogy: to bowl a *googly* involves imparting a right-handed spin to a ball with an action which would appear to be imparting a left-handed spin). The idea is to find a way of using twistor functions of homogeneity -6 to generate a new kind of deformation of twistor space, which can be applied in conjunction with our earlier deformations generated by those of homogeneity $+2$. The present status of this is that there does indeed appear to be a kind of deformation generated by twistor functions of homogeneity -6. These affect the structure of the region of twistor space "at infinity." The twistor interpretation of space-time points needs to be different from the holomorphic curve interpretation considered before, and the details of this remain to be worked out. it is not yet clear whether self-dual solutions of Einstein's vacuum equations can be obtained in this way, but it is quite plausible that they can. The question of how a combined construction might yield full solutions of Einstein's equations is less clear, at the moment.

A non-linear version of the anti-self-dual spin 1 case (left-handed photon, or homogeneity zero twistor function) has been given by R.S. Ward. By deforming complex vector bundles over (part of) $\mathbf{PT}$, one may obtain the general solution of the anti-self-dual Yang-Mills equations (Ward 1977; see also Penrose and Rindler 1986). One may also express the way in which the electromagnetic field affects other (charged) fields using the same idea. For the anti-self-dual electromagnetic field, the vector field of infinitesimal deformation may be taken to be

$$f_{ij}Z^{\alpha}\frac{\partial}{\partial Z^{\alpha}}$$

which affects the structure of $\mathbf{T}$ but not of $\mathbf{PT}$. The "googly problem" for the spin 1 case (helicity $+1$) is in a somewhat similar state to that in the case for spin 2, although rather less well developed.

Many other results and ideas have emerged from twistor theory. For example, one can use solutions of the twistor equation to lower the spin of a massless field by half a unit:

$$\phi_{AB\ldots D} \mapsto \omega^{A}(X)\phi_{AB\ldots D}.$$

 Elements of twistor theory. (Based on a Lecture Delivered in Chittagong 1986, Bangladesh). *Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong.* Ed. J. N. Islam, 1986

The twistor equation on ω^A is precisely what is needed for preservation of the free-field equation when the spin is reduced. This procedure has proved useful in general relativity whereone can apply the procedure in two steps to the (vacuum) curvature tensor to obtain a two-index skew tensor H_{ab} formally resembling a Maxwell field. However the twistor equation does not generally have non-trivial solutions in curved space-time and, instead, one considers only that part of the twistor equation which involves derivatives tangential to some spacelike 2-surface S. Assuming that the topology of S is that of a 2-sphere we find (in general) that this tangential twistor equation has precisely four linearly independent solutions. These solutions therefore define a complex four-dimensional vector space, which is referred to as the *2-surface twistor space* $\mathbf{T}(S)$.

The value of this is that we can perform the standard charge integral on H_{ab} to find the total "charge" surrounded by S. There are ten independent such integrals that one can perform, corresponding to the two possible choices (symmetrically) of element of $\mathbf{T}(S)$ which are used for lowering the spin. These correspond to the ten components of energy, momentum and angular momentum for the total source-plus-field contributions surrounded by S. One anticipates that there is a Hermiticity condition which ensures that these ten independent integrals define ten *real* rather than complex values for these physical quantities.

Many encouraging results have been obtained, showing that in a good many cases this procedure yields physically sensible results. It has emerged, however, that in the general case a modification of this procedure is necessary involving a certain determinant factor. (See Penrose 1982, Tod 1983, Shaw 1983, Huggett and Tod 1985, Penrose and Rindler 1986, Kelly, Tod, and Woodhouse 1986, for further details.)

ACKNOWLEDGEMENTS

I am grateful to the Royal Society of London and the University of Chittagong for making the trip to the conference in Chittagong possible; also to the Institute for The-

oretical Physics at Santa Barbara. Work was partially supported under NSF Grant No. PHY82-17853, supplemented by NASA.

REFERENCES

Hodges, A.P. (1982) *Physica* **114A**, 157-75.

Hodges, A.P. (1985) *Proc. Roy. Soc. London* **A397**, 341-74, 375-96.

Hodges, A.P. and Huggett, S.A. (1980) *Surveys in High Energy Physics*, **1**, 333-53.

Huggett, S.A. and Tod, K.P. (1985) *Introduction to Twistor Theory* (Cambridge University Press, Cambridge).

Hughston, L.P. (1979) *Twistors and Particles*, Lecture notes in Physics No. 97 (Springer-Verlag, Berlin).

Kelly, R.M., Tod, K.P. and Woodhouse, N.M.J. (1986) *Quasi-Local Mass for Small Surfaces* Preprint (to appear).

Penrose, R. (1976) *Gen. Rel. Grav.* **7**, 31-52, 171-6.

Penrose, R. (1982) *Proc. Roy. Soc. London* **A381**, 53-63.

Penrose, R. and Rindler, W. (1984, 1986) *Spinors and Space-Time* Vols. 1 and 2, (Cambridge University Press, Cambridge).

Penrose, R. and Ward, R.S. (1980), in *General Relativity and Gravitation, One Hundred Years after the Birth of Albert Einstein*, ed. A. Held (Plenum, New York).

Perjés, Z. (1977) *Rept. Math. Phys.* **12**, 193-211.

Perjés, Z. (1982) in *Twistor Geometry and Non-Linear Systems* ed. H.D. Doebner and T.D. Palev (Springer-Verlag, Berlin)

Shaw, W.T. (1983) *Proc. Roy. Soc. London* **A390**, 191-215.

428 Elements of twistor theory. (Based on a Lecture Delivered in Chittagong 1986, Bangladesh). *Proceedings of the International Conference on Mathematical Physics organized by the University of Chittagong.* Ed. J. N. Islam, 1986

Tod, K.P. (1983) *Proc. Roy. Soc. London* **A388**, 457-77.
Ward, R.S. (1977) *Phys. Lett.* **61A**, 81-2.

160
ENERGY AND ITS DEFINITION IN GENERAL RELATIVITY

Twelfth Texas Symposium on Relativistic Astrophysics. Eds M. Livio and G. Shavio, 136–46. *Annals of NY Acad. Sci.,* Vol. 470, 1986

We can get some of idea of the then status of quasi-local mass from this survey. It should be mentioned that, 2 lines up from the bottom of p. [145] (in the footnote) the word 'termingly' should be 'seemingly' and that "property K" turned out to be false for the Schwarzschild solution. At the time of writing (May 2010), no clear way of dealing with contorted surfaces has emerged.

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics.* Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

PART IV. GENERAL RELATIVITY

Energy and Its Definition in General Relativity

ROGER PENROSE
Mathematical Institute
University of Oxford
Oxford OX1 3LB, England

BACKGROUND

There is an essential difficulty—and it has been with us since Einstein first put forward his general theory of relativity—concerning the definition, within that theory, of the fundamental concept: energy. The difficulty is present also for the related concepts of momentum and angular momentum, but I shall phrase most of my preliminary discussion in terms of energy (or, equivalently, mass). This difficulty arises because in Einstein's field equation,

$$R_{ab} - \frac{1}{2} R g_{ab} + \lambda g_{ab} = -8\pi G T_{ab} \tag{1}$$

(with a cosmological constant, λ, being allowed for), the energy tensor, T_{ab}, refers to the energy of matter fields only and not to the energy of the gravitational field itself. It is clear on physical grounds that the gravitational field must in fact contribute to the total mass-energy of a physical system. There are two particular ways that this can come about. First, there is the (negative) Newtonian potential energy contribution to the total energy of two (or more) masses. The closer the masses are to each other, the greater will be the amount by which the total mass-energy falls short of the sum of their individual mass-energies. Second, there is the (positive) contribution to the total mass-energy from gravitational waves. These are "ripples" in empty space, so $T_{ab} = 0$ there. Nevertheless, the positivity of gravitational wave energy is a clear implication of the Bondi-Sachs[1] expression for the mass measured at null infinity ($\mathcal{I}^+$).

In mathematical terms, the reason that energy can arise other than from direct contributions from T_{ab} is related to the fact that the covariant divergence law,

$$\nabla_a T^{ab} = 0, \tag{2}$$

does not give rise to an integral conservation law (essentially because of the extra index, "b"), which is in contrast with the case of charge conservation, where $\nabla_a J^a = 0$ does provide us with an integral conservation law for charge. The total mass-energy concept thus must involve contributions not simply arising from "adding up" T_{ab}. Einstein's original proposal for coping with this problem was, of course, to add a "pseudotensor" quantity to T_{ab} that was supposed to represent the gravitational contribution to the energy. However, this quantity has no local coordinate-independent meaning and Einstein regarded it as providing merely a means to calculating the total energy of a system. Even this could be achieved only when the coordinates were specialized to have a very particular form at infinity. Locally, the pseudotensor can

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

always be reduced to zero, emphasizing the fact that the gravitational energy concept is necessarily a nonlocal one.

Physicists seem to have gradually become accustomed to the idea that a meaningful measure of energy for a physical system that incorporates both the gravitational and matter contributions would be too much to ask for, except in the case of the total energy of an entire isolated physical system. For this total energy (for an asymptotically flat space-time), successful expressions do indeed exist: at spatial infinity (i^0), there is the ADM expression[2] and at null infinity ($\mathcal{I}^+$), the Bondi-Sachs expression[1]—the latter having the advantage that by examining its value at different retarded times, the contribution from gravitational or electromagnetic waves can be separated out. These expressions apply to the standard Einstein theory with $\lambda = 0$. With negative λ, for asymptotically anti–de Sitter space-times, a corresponding expression has been introduced by Abbot and Deser,[3] which has been somewhat clarified by Ashtekar and Magnon.[4]

QUASI-LOCAL EXPRESSIONS

It would clearly be physically useful to be able to say something more detailed about the distribution of energy in a physical space-time. In fact, on various occasions, expressions applicable at a more local level have been suggested. I shall refer to an expression that purports to measure the energy (or other physical quantities) surrounded (i.e., "linked") by a closed 2-surface $\mathcal{S}$ as a *quasi-local* expression if it involves only the geometry (or other gauge invariant physical fields) and only in the immediate neighborhood of $\mathcal{S}$. I shall indicate various suggestions that have been made for expressions of this kind.

Historically, the most important attempt at a quasi-local energy expression is that due to Komar.[5] Here one assumes that the space-time, $\mathcal{M}$, is stationary, i.e., it contains a timelike Killing vector, k^a. We assume that k^a is normalized to be a unit vector at infinity, generating the standard time-translations. The suggested energy expression is then

$$(8\pi G)^{-1} \oint \epsilon_{abpq} \nabla^a k^b \, dx^p \wedge dx^q \tag{3}$$

(where ϵ_{abpq} is the alternating tensor). This has the significant property that its value is unchanged as $\mathcal{S}$ is moved continuously through a region of vacuum ($T_{ab} = 0$). Moreover, if $\mathcal{S}$ is taken to infinity, expression (3) agrees with the ADM value (at i^0) and the Bondi-Sachs value (at $\mathcal{I}^+$). However, it has the drawback that the "source-term" for expression (3) is, in effect, the Ricci tensor rather than its trace-reverse—which it should be, according to equation (1) (with $\lambda = 0$)—and this is related to the fact that certain anomalies concerning factors of two tend to arise. (One of these concerns the angular momentum and the other concerns the contributions when matter fields are present, which I shall remark upon shortly.) One further slight drawback is that expression (3) is not, strictly speaking, a completely quasi-local expression since a normalization for k^a at infinity is required.

Winicour and Tamburino[6] adapted Komar's expression to define what they called an energy "linkage," for which k^a need not be assumed actually to be a Killing vector. For the integral to have physical significance, it is necessary to place suitable

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics.* Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

restrictions on k^a. When this is done at $\mathcal{I}^+$, the Bondi-Sachs definition can be recovered. However, in the general case, considerable ambiguities remain.

An expression due to Hawking,[7]

$$(A/16\pi G^2)^{1/2}\,(1 - \frac{1}{2\pi}\oint \rho\,\rho'\,d\mathcal{S}), \tag{4}$$

is also worthy of note. Here, A is the area of the 2-surface $\mathcal{S}$, assumed spacelike and with the topology of a sphere, and ρ and ρ' are the spin-coefficients[8,9] describing the convergence of the null normals to $\mathcal{S}$. When $\mathcal{S}$ is taken to $\mathcal{I}^+$ in a suitable way, the Hawking expression [expression (4)] also reproduces the Bondi-Sachs definition. However, in a general way, expression (4) cannot be used just as it stands and it must be supplemented by some condition such as of "maximal roundness." (Otherwise unreasonable answers would arise, notably nonzero values for flat space-time.)

A different definition has been more recently proposed by Ludvigsen and Vickers.[10] This arose out of their work developing Witten's argument for the positivity of the ADM mass, which was done in order to provide one of the proofs of the somewhat stronger result that the Bondi-Sachs mass is likewise positive-definite. The link with this positivity theorem gives their approach some physical plausibility, but their definition is not really a quasi-local one in the sense described above since it depends upon information carried in from infinity on a null hypersurface through $\mathcal{S}$.

Finally, there is a quasi-local definition that I myself proposed a few years ago[11,12] for the mass, momentum, and angular momentum, surrounded by a spacelike topological 2-sphere. The remainder of this account will be concerned with this expression: its motivation, definition, and latest results.

MOTIVATION

The paradigm for a quasi-local expression would appear to be provided by Maxwell theory. Here, one has the expressions,

$$Q = \frac{1}{8\pi}\oint_{\mathcal{S}} {}^*F_{ab}\,dx^a \wedge dx^b \tag{5}$$

and

$$0 = \frac{1}{8\pi}\oint_{\mathcal{S}} F_{ab}\,dx^a \wedge dx^b, \tag{6}$$

respectively, for the electric and magnetic charges surrounded by $\mathcal{S}$. Here

$${}^*F_{ab} = \frac{1}{2}\,e_{abcd}\,F^{cd} \tag{7}$$

and the Maxwell equations,

$$\nabla_{[a}F_{bc]} = 0,\quad \nabla^a F_{ab} = 4\pi J_b, \tag{8}$$

are being assumed. Gauss's law, generalized to a general-relativistic setting, enables

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

equation (5) to be converted to a 3-volume integral,

$$Q = \frac{1}{6}\int_{\Sigma} J^a \epsilon_{apqr}\, dx^p \wedge dx^q \wedge dx^r, \tag{9}$$

of charge density over the compact 3-region, Σ, where $\mathcal{S} = \partial\Sigma$.

It should be pointed out that the mere existence of a "quasi-local" expression for charge, such as equation (5), in itself implies a weak form of charge conservation. For it makes no difference to the value of Q how the 2-surface $\mathcal{S}$ is spanned. Any two compact 3-regions, Σ, Σ', whose boundaries are the same surface,

$$\partial\Sigma = \mathcal{S} = \partial\Sigma', \tag{10}$$

would be deemed to intercept the same total charge, namely Q, even in the absence of some explicit formula such as equation (9). One needs only some convincing reason to believe in the physical appropriateness of the proposed quasi-local expression [here equation (5)]. Its conservation, in this weak sense, is a necessary consequence.

There is, however, a stronger form of conservation that does not follow from the mere existence of a quasi-local expression. Strong conservation arises when the "charge" value remains constant whenever the 2-surface $\mathcal{S}$ is moved continuously through a source-free region. It is clear that this requirement of strong conservation is indeed satisfied in Maxwell theory. In physical terms, we say that the field is itself "uncharged." We do not, however, expect that a corresponding property would hold for Yang-Mills theory—nor, indeed, for general relativity since, as was remarked above, the gravitational field can itself physically possess mass-energy, albeit in some nonlocal form. The existence of a quasi-local energy expression in general relativity is perfectly consistent with these physical requirements.

To motivate my own suggestion for such a quasi-local expression, consider first the linearized theory. In this case, we do expect a strong conservation law, with the energy of gravitation being quadratic in the field (as can be seen, for example, in the Bondi-Sachs flux at $\mathcal{J}^+$). Thus, it can be ignored in the linearized limit. The idea is to take the (weak field) curvature tensor, R_{abcd}, which is a spin-2 object, and reduce it to a spin-1 object—which behaves formally like a Maxwell field—by contracting off two of its indices with a suitably defined object, W^{ab} ($= -W^{ba}$):

$$R_{abcd}\, W^{cd}. \tag{11}$$

The equation to be satisfied by W^{ab} turns out to be

$$\nabla^{(a} W^{b)c} - \nabla^{(a} W^{c)b} + g^{a[b} \nabla_d W^{c]d} = 0 \tag{12}$$

(round parentheses denoting symmetrization and square brackets denoting antisymmetrization). Putting

$$k^a = \frac{2}{3}\, \epsilon^{abcd} \nabla_b W_{cd}, \tag{13}$$

we find that k^a is automatically a Killing vector of the flat background space-time of the linearized theory. Computing the charge integral for the "Maxwell field" in equation (11), we actually obtain[11,12] the energy, momentum, or angular momentum component of the sources surrounded by $\mathcal{S}$, which correspond to the Killing vector

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

defined by equation (13). (The results agree with those of Sachs and Bergmann,[13] though their procedures were completely different.)

As stated, the reason this works is somewhat obscure. Things are rather more transparent if a spinor formalism is used. Using the abstract index notation,[9] we have, in the vacuum region,

$$R_{abcd} = \Psi_{ABCD}\,\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\bar{\Psi}_{A'B'C'D'}, \tag{14}$$

where Ψ_{ABCD} is symmetric and satisfies

$$\nabla^{AA'}\Psi_{ABCD} = 0 \tag{15}$$

(from the linearized Bianchi identity); we can also set

$$W^{ab} = \omega^{AB}\epsilon^{A'B'} + \epsilon^{AB}\bar{\omega}^{A'B'}, \tag{16}$$

where ω^{AB} is symmetric and satisfies

$$\nabla^{(A}_{A'}\omega^{BC)} = 0 \tag{17}$$

[which is the spinor form of equation (12)]. We now find that equation (11) has the spinor expression,

$$\chi_{AB}\,\epsilon_{A'B'} + \epsilon_{AB}\bar{\chi}_{A'B'}, \tag{18}$$

where

$$\chi_{AB} = \psi_{ABCD}\omega^{CD} \tag{19}$$

is symmetric and satisfies the spinor form of Maxwell's free-space equation,[9]

$$\nabla^{AA'}\chi_{AB} = 0. \tag{20}$$

This equation is an immediate consequence of equations (15) and (17). Indeed, equations (15) and (20) are particular cases of the massless free-field equation for spin $n/2$:

$$\nabla^{AA'}\phi_{AB\ldots L} = 0, \tag{21}$$

where $\phi_{AB\ldots L}$ has n symmetric indices. As a consequence of equations (17) and (21), the quantity,

$$\phi_{AB\ldots JKL}\omega^{KL}, \tag{22}$$

which is symmetric with $n - 2$ indices, satisfies the massless free-field equation for spin $(n - 2)/2$. Thus, the quantity, ω^{AB}, subject to equation (17), provides a procedure in the general case of an arbitrary massless field for lowering the spin by one unit.

In Minkowski space-time, equation (12) has twenty linearly independent read solutions [i.e., equation (17) has ten independent complex solutions]. However, for the independent real solutions, the Killing vector of equation (13) vanishes and the integrals of equation (11) must therefore also vanish. The remaining ten independent real solutions of equation (12) provide integrals that do not generally vanish and they give the ten components of energy, momentum, and angular momentum for the sources surrounded by $\mathcal{S}$. Indeed, one obtains a strong conservation law for these quantities

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

In general relativity proper, however, equation **(12)** [or, equivalently, equation **(17)**] will in general have no solutions (other than zero). This is not surprising since equation **(12)** represents sixteen independent equations on only six independent unknowns [or, in complex terms, equation **(17)** represents eight equations on three unknowns]. Indeed, it would be physically undesirable for equation **(12)** to have nontrivial solutions in the general case since these would lead to strong conversation laws in general relativity for the corresponding energy, momentum, or angular momentum components. All we seek (at this stage) is a reasonable quasi-local *definition* of such quantities so that weak, rather than strong, conservation will hold. We require something that "looks like" W^{ab} (or ω^{AB}) as far as $\mathcal{S}$ is concerned, but which need not extend into the ambient space-time, $\mathcal{M}$, in any specific way. It is natural to demand, therefore, merely that the "tangential part" of equation **(12)** [or **(17)**] be satisfied. However, we now find that the resulting equations are underdetermined since there are only four equations for the six unknowns (or two equations for three unknowns, in the complex description).

The way around this impasse turns out to be to lower the spin in two steps of one-half each (rather than in a single step). This is achieved, in flat space, by use of a spinor quantity, ω^A, which satisfies

$$\nabla_{A'}^{(A}\omega^{B)} = 0. \tag{23}$$

For, by the same reasoning as before, if $\phi_{AB\ldots L}$ (symmetric) satisfies the massless free-field equation **(21)** for spin $n/2$, then

$$\phi_{AB\ldots KL}\omega^{L} \tag{24}$$

satisfies the same equation for spin $(n-1)/2$. Equation **(23)** is called the twistor equation[14,15] and its solutions, in Minkowski space-time, constitute a four-complex-dimensional vector space, T, referred to as twistor space. (This is just one of many possible definitions of twistor space. It suffices for our purposes here. The conformal supersymmetry generators of Wess and Zumino[16] are also effectively twistors in this sense.)

Again, equation **(23)** does not generally have (nontrivial) solutions in a curved space-time, $\mathcal{M}$, since there are six complex equations for two complex unknowns. However, restricted to $\mathcal{S}$, we now obtain two equations for two unknowns. Elementary index theory shows us that indeed there is, at least in the generic case, a four-complex-dimensional space of solutions of these tangential equations, provided that $\mathcal{S}$ is spacelike and topologically a 2-sphere. This space is called the 2-surface twistor space and is denoted by $\mathrm{T}(\mathcal{S})$. The required analogues of ω^{AB}, needed for defining the analogues of W^{ab} used in equation **(11)**, are now the elements of the symmetric tensor product,

$$\mathrm{T}(\mathcal{S}) \odot \mathrm{T}(\mathcal{S}). \tag{25}$$

It is remarkable that these objects do not seem to be definable directly, and it appears to be necessary to employ "spin-½" objects in order to define what would appear to be essentially "integral spin" concepts, namely, energy-momentum and angular momentum.

 Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

TECHNICAL DEFINITIONS

The tangential part of equation (23) is conveniently expressed in terms of the compacted spin-coefficient formalism.[9,17] The equations become

$$\eth'\omega^0 = \sigma'\omega^1, \quad \eth\omega^1 = \sigma\omega^0, \tag{26}$$

where $\eth'$ and $\eth$ are suitably defined (covariant) derivative operators within $\mathcal{S}$, with the spin-frame being chosen to have its "flagpole" directions orthogonal to $\mathcal{S}$. A solution of equation (26) on $\mathcal{S}$ [element of $\mathsf{T}(\mathcal{S})$] may be denoted by Z^α (a 2-surface twistor), where the four-dimensional abstract index, α, refers to the 4-space $\mathsf{T}(\mathcal{S})$. Taking Z^α and $\tilde{Z}^\alpha$ to be two arbitrary elements of $\mathsf{T}(\mathcal{S})$ [given, respectively, by the solutions, ω^0, ω^1 and $\tilde{\omega}^0$, $\tilde{\omega}^1$, of equation (26)], we can tentatively define the element, $A_{\alpha\beta}$, of $\{\mathsf{T}(\mathcal{S}) \odot \mathsf{T}(\mathcal{S})\}^*$ by the relation,

$$A_{\alpha\beta} Z^\alpha \tilde{Z}^\beta = (4\pi i G)^{-1} \oint_{\mathcal{S}} (P\omega^0\tilde{\omega}^0 + Q(\omega^0\tilde{\omega}^1 + \omega^1\tilde{\omega}^0) + R\omega^1\tilde{\omega}^1) d\mathcal{S}. \tag{27}$$

Here

$$P = \Psi_1 - \Phi_{0,1} = R_{abcd} l^a m^b \bar{m}^c m^d, \tag{28}$$

$$Q = \Psi_2 - \Phi_{1,1} - \Lambda + \lambda/6 = \frac{1}{2} R_{abcd} (l^a n^b + \bar{m}^a m^b) \bar{m}^c m^d + \lambda/6, \tag{29}$$

and

$$R = \Psi_3 - \Phi_{2,1} = R_{abcd} \bar{m}^a n^b \bar{m}^c m^d \tag{30}$$

[with a cosmological constant, λ, being allowed for, as in equation (1)]. The null tetrad $(l^a, m^a, \bar{m}^a, n^a)$ is that defined by the spin-frame.

This (tentatively) defines the total angular momentum twistor, $A_{\alpha\beta}$, for the matter + gravitation surrounded by $\mathcal{S}$. In linearized theory (flat space), this takes the form,[15]

$$A_{\alpha\beta} = \left[\begin{array}{c|c} 0 & \begin{array}{c}\text{energy-}\\ \text{momentum}\end{array} \\ \hline \begin{array}{c}\text{energy-}\\ \text{momentum}\end{array} & \begin{array}{c}\text{angular}\\ \text{momentum}\end{array} \end{array}\right], \tag{31}$$

and a Hermitian (squared) norm, $\|Z^\alpha\|$, of signature, $(+\,+\,-\,-)$, is defined on T, whose natural extension to $\{\mathsf{T} \odot \mathsf{T}\}^*$ allows us to define the squared rest-mass as

$$m^2 = -\frac{1}{2} \|A_{\alpha\beta}\|. \tag{32}$$

The particular form of equation (31) entails that only ten of the possible twenty independent components of $A_{\alpha\beta}$ survive in the linearized theory (as noted earlier). It is not yet known whether an analogous reduction occurs in the full theory in all cases Moreover, there is, as yet, no clear agreement as to the best definition of the Hermitian norm on $\mathsf{T}(\mathcal{S})$ in all cases, although in many circumstances the choice is clear.

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics.* Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

SPECIFIC EXAMPLES

At null infinity, $\mathcal{I}^+$, this definition provides an energy-momentum 4-vector that is in complete agreement[11] with that of Bondi-Sachs. It also provides a definition of angular momentum that has some significant advantages[15] over those that had been suggested earlier,[18] notably proper agreement with linearized theory (for "bad cuts" of $\mathcal{I}^+$). At spacelike infinity, i^0, agreement is obtained with the ADM energy-momentum[20] and the Ashtekar-Hansen angular momentum.[20] (There seems also to be agreement with the Abbot-Deser-Ashtekar-Magnon definition for the negative cosmological constant.[21])

Several striking results have been obtained by Tod[22] for finite surfaces $\mathcal{S}$. For example, if $\mathcal{S}$ is drawn arbitrarily in a constant-time hypersurface in the Schwarzschild solution (or, indeed, in any rotationally symmetric hypersurface, $t - f(r)$, in standard coordinates), then the Schwarzschild mass is obtained for the m of equation (32) whenever $\mathcal{S}$ links the source (once) and the value of $m = 0$ is obtained if $\mathcal{S}$ does not link the source. For the Reissner-Nordstrom solution, there are somewhat analogous results, but now there is the (expected) positive contribution from the energy density of the electromagnetic field. Thus, the value of m depends on its particular location in addition to whether or not it links the source. [It is worth remarking here that if the Komar expression in expression (3) is used, then the contribution from the electromagnetic energy density comes out wrong by a factor of two.]

Surfaces $\mathcal{S}$ drawn in $t = f(r)$ hypersurfaces in these solutions have the special property that they can be embedded in conformally flat space-time without changing their first or second fundamental forms. Such surfaces we refer to as *uncontorted:* they are *contorted* if they *cannot* be so embedded. It is natural to conjecture that for a general (contorted) $\mathcal{S}$ in the Schwarzschild solution, it should continue to be the case that the m obtained is the Schwarzschild mass when $\mathcal{S}$ links the source (once) and zero when $\mathcal{S}$ does not link the source. We shall see the status of this conjecture shortly.

Tod[22] also examined Friedmann-Robertson-Walker models and found, in particular, that for a spherically symmetrical surface, m has the same value as would be the case in Euclidean space with a sphere of the same area that was immersed in a fluid of the same density as that of the model. Thus, for example, in the closed models, $\mathcal{S}$ reaches its maximum for an "equatorial" 2-sphere and reduces back to zero when the entire universe is encompassed. One can also say quite generally, with this definition, that the total mass-energy of any closed universe is zero (in agreement with expectations). Moreover, in any such universe, the mass m on one side of any $\mathcal{S}$ is always equal to that on the other.

One of the most striking of Tod's[22] results concerns 2-surfaces $\mathcal{S}$ in a conformally flat 3-surface $\mathcal{H}$ of time-symmetry. If $\mathcal{S}$ is continuously deformed throughout a vacuum region of $\mathcal{H}$, then it turns out that m does not change. Furthermore, one may envisage $\mathcal{S}$ linking one or the other of two source regions (yielding the results m_1, m_2, respectively) or else linking both sources together (yielding the result $m_{1,2}$). Tod finds the physically expected inequality,

$$m_{1,2} < m_1 + m_2, \tag{33}$$

and he also finds the result that for large separations, the difference agrees with the correct Newtonian potential-energy contribution. This result is particularly striking

Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

because potential-energy effects did not directly enter into the motivational considerations for the definition of m given here.

A more recent result obtained by Tod[23] concerns a static "black hole in a cage." The cage is not to prevent the hole from "escaping," but merely to distort it from spherical shape. For given surface area (which would not be changed if the cage is altered "adiabatically"), one would anticipate that the hole configuration of least energy would be that of spherical shape. Thus, indeed, Tod finds the inequality,

$$m^2 \geq A/16\pi G^2, \tag{34}$$

with equality only in the spherically symmetrical Schwarzschild case.

THE η-FACTOR MODIFICATION

These results are all very gratifying, but difficulties have arisen from another calculation performed by Tod[24] and checked by R. M. Kelly. The light cone, $\mathscr{C}$, of some arbitrary point, p, in a space-time, $\mathscr{M}$, is considered. A timelike vector, t^a, is selected at p and cross sections, $\mathscr{S}_r$, of constant affine distance, r, from p are drawn, where the various affine parameters along the rays of $\mathscr{C}$ are normalized against t^a at p. The question is: what is the value of m (and the structure of $A_{\alpha\beta}$) for $\mathscr{S}_r$ at small r? When $\mathscr{M}$ is not vacuum at p, the expected energy-momentum,

$$\frac{4}{3}\pi r^3 T_{ab}t^b, \tag{35}$$

is indeed found at order, r^3 (as could be inferred in any case from agreement with the linearized limit). When T_{ab} vanishes, the first possibility for a contribution occurs at order, r^5. It is found that there is such a contribution for the definition in equation (27) and that if $\mathscr{M}$ is the Schwarzschild solution, then this contribution is nonvanishing when t^a does not lie in a (t, r)-plane. This contradicts the aforementioned conjecture, as it is stated, but it is worse than that: the m^2 obtained turns out to be negative (i.e., m is imaginary). These are cases when the surfaces $\mathscr{S}_r$ are contorted.

In the meantime, I had by myself obtained a procedure from attempting to prove the opposite result, that, in a suitable sense, this conjecture ought to be true. I was able to show that if a certain "Property K" were true for the Schwarzschild solution, then any $\mathscr{S}$ not linking the source, whether contorted or not, should yield zero for m (and for $A_{\alpha\beta}$) if a certain modification is made to the integral in equation (27). This consists of including a factor, η, that multiplies the entire integrand, where η is proportional to the determinant,

$$\begin{vmatrix} \omega_1^0 & \omega_2^0 & \omega_3^0 & \omega_4^0 \\ \omega_1^1 & \omega_2^1 & \omega_3^1 & \omega_4^1 \\ \eth\omega_1^0 & \eth\omega_2^0 & \eth\omega_3^0 & \omega_4^0 \\ \eth'\omega_1^1 & \eth'\omega_2^1 & \eth'\omega_3^1 & \eth'\omega_4^1 \end{vmatrix}, \tag{36}$$

with ω_1^A, ω_2^A, ω_3^A, and ω_4^A providing four independent solutions of equation (26). When $\mathscr{S}$

is uncontorted (and in certain other cases, such as when $\mathcal{S}$ is an arbitrary "cut" of $\mathcal{I}^+$), expression (36) is constant over $\mathcal{S}$ and so η can be chosen to be unity. In the general case, we need some normalization for η, with one suggested possibility being

$$\oint_{\mathcal{S}} \log \eta \, d\mathcal{S} = 0. \tag{37}$$

When this modification is made to equation (27), Tod and Kelly find that m now vanishes to order, r^5. However, Property K itself remains unestablished. This property asserts that the Killing spinor[15,25] [which is a particular solution of equation (17) arising in any type {22}—i.e., type D—space-time] always belongs to $\mathsf{T}(\mathcal{S}) \odot \mathsf{T}(\mathcal{S})$. This property must be false for the C-metrics[26] and for the Kerr solution,[26] but the circumstances for the Schwarzschild solution seem to be considerably more favorable.*

Shaw[27] has recently performed calculations of a reverse nature to those of Tod and Kelly, where expansions in terms of negative powers of an affine parameter, r, are obtained ("large spheres"). Physically reasonable results are so far obtained for the definition of equation (27) (with or without η-factor), though the Hawking and Ludvigsen-Vickers definitions contain some anomalies.

REFERENCES

1. Bondi, H. 1960. Nature **186:** 535; Bondi, H., M. G. J. van der Burg & A. W. K. Metzner. 1962. Proc. R. Soc. London **A269:** 21; Sachs, R. K. 1962. Proc. R. Soc. London **A270:** 103.
2. Arnowitt, R., S. Deser & C. W. Misner. 1962. *In* Gravitation: An Introduction to Current Research. L. Witten, Ed. Wiley. New York.
3. Abbot, L. F. & S. Deser. 1982. Nucl. Phys. **B195:** 76.
4. Ashtekar, A. & A. Magnon. 1984. Class. Quantum Grav. **1:** L30.
5. Komar, A. 1959. Phys. Rev. **113:** 934.
6. Tamburino, L. & J. Winicour. 1966. Phys. Rev. **150:** 1039.
7. Hawking, S. W. 1968. J. Math. Phys. **9:** 598.
8. Newman, E. T. & R. Penrose. 1962. J. Math. Phys. **3:** 566; 1963. J. Math. Phys. **4:** 998.
9. Penrose, R. & W. Rindler. 1984. Spinors and Space-Time, vol. 1: Two-Spinor Calculus and Relativistic Fields. Cambridge University Press. Cambridge.
10. Ludvigsen, M. & J. A. G. Vickers. 1982. Bondi Momentum and its Quasi-Local Null Surface Extension. University of Canterbury Preprint. Christchurch, New Zealand.
11. Penrose, R. 1982. Proc. R. Soc. London **A381:** 53.
12. Penrose, R. 1984. J. C. Maxwell, the Sesquicentennial Symposium. M.S. Berger, Ed. Elsevier. Amsterdam/New York.
13. Sachs, R. K. & P. G. Bergmann. 1958. Phys. Rev. **112:** 674.
14. Penrose, R. 1967. J. Math. Phys. **8:** 345.
15. Penrose, R. & W. Rindler. 1986. Spinors and Space-Time, vol. 2: Spinor and Twistor Methods in Space-Time Geometry. Cambridge University Press. Cambridge.
16. Wess, J. & B. Zumino. 1974. Nucl. Phys. **70:** 39.
17. Geroch, R., A. Held & R. Penrose. 1973. J. Math. Phys. **14:** 874.
18. Winicour, J. 1980. *In* General Relativity and Gravitation, vol. 2. A. Held, Ed. Plenum. New York.

*After this lecture was given, some results of N.M.J. Woodhouse,[28] using an original calculational procedure, have confirmed the Tod-Kelly "small spheres" calculation, but also seemingly yield a new anomaly, not removed by the η-factor, for "small ellipsoids." **[Note added in proof:** This new anomaly has subsequently disappeared.]

 Energy and its definition in general relativity. *Twelfth Texas Symposium on Relativistic Astrophysics*. Eds M. Livio and G. Shavio, 136–46. Annals of NY Acad. Sci., Vol. 470, 1986

19. ASHTEKAR, A. & R. O. HANSEN. 1978. J. Math. Phys. **19:** 1542.
20. SHAW, W. T. 1983. Proc. R. Soc. London **A390:** 191.
21. KELLY, R. M. Private communication.
22. TOD, K. P. 1983. Proc. R. Soc. London **A388:** 457.
23. TOD, K. P., R. M. KELLY & N. M. J. WOODHOUSE. To appear.
24. TOD, K. P. 1984. Twistor Newsletter **18**. Math. Inst. Oxford.
25. WALKER, M. & R. PENROSE. 1970. Commun. Math. Phys. **18:** 265.
26. KRAMER, D., H. STEPHANI, M. MACCALLUM & E. HERLT. 1980. Exact Solutions of Einstein's Field Equations. Deut. Verlag Wissenschaften. Berlin.
27. SHAW, W. T. 1985. Twistor Newsletter **19**. Math. Inst. Oxford.
28. WOODHOUSE, N. M. J. 1985. Twistor Newsletter **19**. Math. Inst. Oxford.

161
TWO-SURFACE TWISTORS AND HYPERSURFACE TWISTORS

Twistor Newsletter **21** (1986). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. Chapman and Hall/CRC, London, 1995

 Two-surface twistors and hypersurface twistors. *Twistor Newsletter* **21** (1986). Also published in *Further Advances in Twistor Theory*, Vol. 2. Eds L. J. Mason and L. Hughston. 1995

§II.4.25 **Two-surface twistors and hypersurface twistors** *by R.Penrose* (**TN** 21, February 1986)

Let $\mathcal{S}$ be a real analytic 2-surface (topology S^2) in an analytic spacelike hypersurface $\mathcal{H}$ in an analytic space-time, and let $\mathbf{C}\mathcal{S}$, $\mathbf{C}\mathcal{H}$ be respective complex 'thickenings' of $\mathcal{S}$, $\mathcal{H}$. Let ω^A be a solution of the 2-surface twistor equation on $\mathcal{S}$. Then

$$\xi_A \xi^B \eta^{C'} \nabla_{BC'} \omega^A = 0 \tag{1}$$

whenever $\xi^A \eta^{A'}$ is tangent to $\mathcal{S}$. Let ζ be a β-curve in $\mathbf{C}\mathcal{H}$ with associated (tangent) spinor λ_A, so that

$$N^{BA'} \lambda_B \lambda^A \nabla_{AA'} \lambda_{C'} = 0 \tag{2}$$

along ζ, where N^a is normal to $\mathcal{H}$. (Note that $N^{BA'} \lambda_B \lambda^A$ is tangent to ζ.) Let $\mathcal{T}^*$ be the dual hypersurface twistor space for $\mathcal{H}$. Consider the function Ω, on a region $\mathcal{R}$ of $\mathcal{T}^*$, defined as the value of

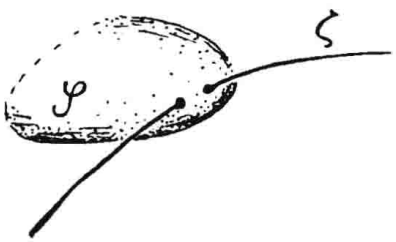
$$\omega^A \lambda_A$$

at the intersection point of ζ with $\mathcal{S}$—assuming for the moment, that this point is unique. It is clear from the holomorphic nature of this construction that Ω is holomorphic on $\mathcal{R}$. Now consider the situations where ζ is *tangent* to $\mathcal{S}$. If Ω is to remain holomorphic we require that the derivative of $\omega^A \lambda_A$ along the β-curve ζ vanishes. Indeed we have

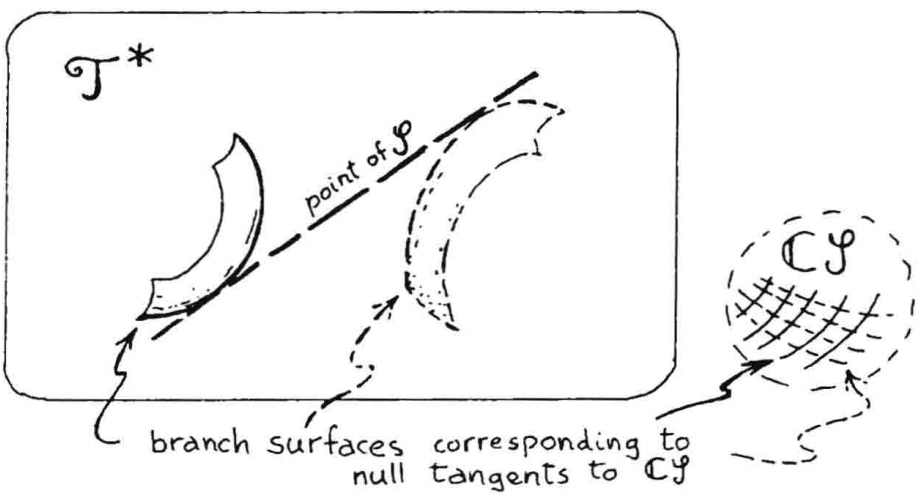
$$N^{BA'} \lambda_B \lambda^A \nabla_{AA'} (\omega^C \lambda_C) = N^{BA'} \lambda_B \lambda^A \lambda_C \nabla_{AA'} \omega^C + \omega^C N^{BA'} \lambda_B \lambda^A \nabla_{AA'} \lambda_C = 0$$

since the first term vanishes by (1) and the second by (2).

Though necessary, the vanishing of this derivative is not sufficient for Ω to be holomorphic at the points of $\mathcal{T}^*$ where ζ touches $\mathcal{Y}$. For near such places ζ will have 2 intersections with $\mathcal{Y}$ and so the function Ω is liable to be 2-valued there. The condition just established ensures that to *first order* these 2 values coincide. But they need not coincide to higher orders.



The (linearly independent) elements of $\mathbf{T}(\mathcal{Y})$ (i.e. 4 independent ω^A's) provide *coordinates* ('linear' functions) on $\mathcal{R}$ via their corresponding functions Ω. Such coordinates are characterized by the fact that the branching vanishes to 1st order—and this fact serves to *define* the 2-surface twistor space $\mathbf{T}(\mathcal{Y})$ in terms of $\mathcal{T}^*$. Only when $\mathcal{T}^*$ is flat (so 3-surface twistors exist on $\mathcal{H}$) can holomorphic coordinates extend single-valuedly to the *whole* of $\mathcal{T}^*$, so only in this case is the branching completely absent.



Similar arguments can be applied to β-planes (in place of β-curves) in the complexified space-time, when these exist.

162
EXPONENTIATING A RELATIVE H^2

Twistor Newsletter **21** (1986). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 Exponentiating a relative H^2. *Twistor Newsletter* **21** (1986). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.15 Exponentiating a Relative H^2 *by R. Penrose* (TN 21, February 1986)

As an approach to solving the googly problem, I suggested in §III.4.11 and §III.4.12 that the required googly information might reside in the way that $\mathbb{I}$ is attached to $\mathbb{T}$, and that an infinitesimal version of this information might be described by an element of $H^2_{\mathbb{I}}(\mathcal{Q}, \widehat{d\Omega}^2)$, such an element being defined by a -6-homogeneity twistor function. Here $\mathcal{Q}$ is some neighbourhood of $\mathbb{I}$ in $\mathbb{T}(= \mathbb{T} \cup \mathbb{I})$ and $\widehat{d\Omega^2}$ is the sheaf of closed holomorphic 3-forms proportional to

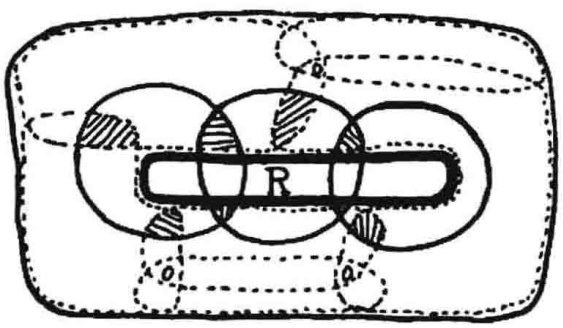
$$\mathcal{D}Z = \frac{1}{6} Z \wedge dZ \wedge dZ \wedge dZ$$

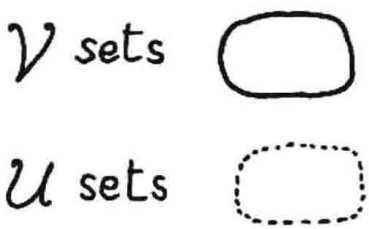
(Recall that $\mathbb{I}$ — 'non-projective blown-up I'- is given by the finite limits of

$$(\omega^A \pi_{B'}, \pi_{A'} \pi_{B'})$$

as $\pi_{A'} \to 0$, $\omega^A \to \infty$). An essential piece of understanding for the success of this proposal would be to see how a relative H^2 can be exponentiated to provide a finite deformation. I shall suggest here a possible procedure for doing this.

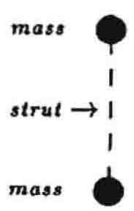
Let M and R ($\subset M$) be complex manifolds (or M a complex manifold and R a closed subset of M). We consider the relative cohomology group $H^2_R(M, \Theta)$ where Θ is the sheaf of holomorphic vector fields on M. A representative cocycle for this has vector fields defined on the various triple intersections of open sets of a covering $\{\mathcal{U}\} \cup \{\mathcal{V}\}$ of M, where $\{\mathcal{U}\}$ covers $M - R$ and $\{\mathcal{V}\}$ covers some (suitably small) neighbourhood of R in M.





Whenever *all* 3 sets belong to $\{\mathcal{U}\}$, the vector field must be *zero*. We wish to deform the way that R is attached to M by exponentiating the vector fields on each triple overlap and to use these as transition functions between the patches. This does not properly give a deformation of M (which would have been achieved by transition functions between *pairs* of open sets). There is always at least one *extra* set $\mathcal{V}_i \in \{\mathcal{V}\}$ involved. Roughly speaking, we get, instead, a deformation of $\mathcal{V}_i$ — so the deformation affects only a neighbourhood of R in M, outside of which one has to revert to the original structure of M.

What kind of space does this result in? My suggestions is that it is a kind of 'generalized orbifold'. To picture the kind of thing that would be involved, recall the Weyl-Levi-Civitá space-time representing two masses propped apart by a strut. As one approaches the strut, one finds that the curvature is finite and that the strut is described merely



by a conical singularity. By altering the period at which the angular coordinate must be identified, the singularity can be locally removed. However, this cannot be done globally, and is to be thought of as a local extension which removes the singularity only near the strut. To attempt to extend the re-identification of the angular coordinate globally would lead to an inconsistency elsewhere (as the masses are surrounded). I am suggesting that the deformed attachment of R to M is somewhat like this space-time with the strut singularity locally 'removed' — though instead of simply having an angular identification to re-adjust, one would anticipate a whole functional freedom available for the re-adjustment — a holomorphic function of 3 complex variables worth. The re-attachment of R would be locally non-singular (like the strut above) but would not be extended consistently to the whole of $M - R$.

Note that for this to be a sensible procedure the space must have a symmetry: axi-symmetry in the case of the Weyl-Levi Civitá space-time. In the case of $\mathbf{T}$, the idea would be to use a symmetry of the fibres in $\begin{array}{c}\mathbf{T}\\ \downarrow \\ \mathbb{P}\mathbf{T}\end{array}$. However, because -6-functions are involved, we are now looking for transformations which have the form $(1+f_{-6})^{1/6}$ along the fibres — the fibre-preserving transformations for which $\lceil\dot{Z}\lceil\dot{Z}\, dZ\wedge dZ\wedge dZ\wedge dZ$ (i.e., $\pi_{A'}\pi_{B'}d^4Z$) is invariant. (Compare my 'googly photon' suggestion §O.3.7.) How to use such a thing in order to single out the correct googly maps is not yet wholly clear. The hope is that it can be tied in with some ways of describing googly maps in terms of *forms*, that Toby Bailey, Michael Singer and I have been developing (see also §III.4.1).

Appreciation to M.A. Singer is expressed.

163
GRAVITATIONAL FIELD MASS

Gravitational collapse and relativity; Yamada Conference XIV. Eds H. Sato and T. Nakamura, 43–59. World Scientific, Singapore, 1986

450 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

GRAVITATIONAL FIELD MASS

by

Roger Penrose

ABSTRACT

The author's definition for the mass-momentum/angular momentum surrounded by a spacelike 2-surface with S^2 topology is presented, this definition being motivated by some ideas from twistor theory in relation to linearized gravitational theory. The status of this definition is examined in relation to many examples which have been worked out (largely by K.P. Tod). The reason for introducing a slight modification to the original definition is also presented.

INTRODUCTION

This is a report on the status of the quasi-local definition that I suggested some years ago[1),2)] for the mass, momentum and angular momentum surrounded by a general spacelike 2-surface $\mathcal{S}$, with S^2 topology, in an arbitrary space-time manifold $\mathcal{M}$. The intention behind the definition is that it should fully take into account the non-local field energy contributions. I shall first outline the basic argument motivating the definition, explaining in detail what the definition actually is. Then I shall describe various situations where the mass, according to this definition, has been computed (mainly by Tod[3),4)]). The results of these calculations are generally encouraging, but it has become clear that a certain modification[5),6)] of the original definition is needed. Happily, all the good results remain unaffected by this modification, whereas certain anomalous ones for small surfaces found by Tod, Kelly and Woodhouse[7)] are eliminated.

Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

GENERAL PROCEDURE

It has been clear for a long time that vacuum gravitational field can contribute to the total mass of a system in a non-local way. From the Bondi-Sachs mass-loss formula one can infer that gravitational waves contribute positively to the mass of a system. However, even at infinity this contribution is non-local. The complex news function N, whose squared modulus measures the gravitational energy flux across future null infinity $\mathscr{I}^+$, is not locally defined at points of $\mathscr{I}^+$, but requires some constant of integration or global scaling for its definition. (See Penrose and Rindler[5] - it is even possible to have $N \neq 0$ throughout an open region of $\mathscr{I}^+$ which has a neighbourhood in $\mathscr{M}$ where the physical metric is completely flat!) In a somewhat less clear way, the mutual gravitational potential energy of a pair of masses contributes non-locally, but now negatively, to the total mass of the system.

Maxwell's electromagnetic theory provides a paradigm, in which a conserved charge exists which can be computed by performing an integral of the dual $*F_{ab}$ (regarded as a 2-form) of the Maxwell field tensor F_{ab} over a 2-surface such as $\mathscr{S}$. The idea is to do a similar thing with the Riemann curvature tensor R_{abcd}, but since this has too many indices, we shall need an appropriate way of "snuffing off" two of them. The motivation for how this is to be done comes from twistor theory and from a desire to obtain the correct results in the weak-field limit of general relativity. What is needed is a means of reducing the spin from 2 (weak-field limit of Einstein's theory) to 1 (Maxwell theory). In fact one way of regarding a twistor is as a spin-lowering operator (in flat space) which lowers the spin by ½. The required spin reduction can then be achieved by applying this procedure twice to the curvature tensor; and the charge-integral expression is then applied to the resulting Maxwell field.

TWISTOR PROCEDURE IN MINKOWSKI SPACE

For the required purpose as described above, we regard a twistor as a solution of the twistor equation

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

$$\nabla^{(A}_{A'}\,\omega^{B)} = 0 \qquad (1)$$

(using standard 2-spinor notation, cf. Penrose and Rindler[8]). The solutions of (1) are precisely the objects which take a solution of the massless free-field equation

$$\nabla^{AA'}\,\phi_{AB\ldots L} = 0 \qquad (2)$$

for spin $\frac{1}{2}n$ (where $\phi_{AB\ldots L} = \phi_{(AB\ldots L)}$ has n indices) to one for spin $\frac{1}{2}(n-1)$, according to

$$\phi_{AB\ldots L} \mapsto \omega^{A}\phi_{AB\ldots L} \ . \qquad (3)$$

We assume, for the momentum that our space-time is Minkowski space $\mathbb{M}$.

The solutions of (1) form a complex four-dimensional vector space $\mathbb{T}$ referred to as twistor space. Each element Z^{α} of $\mathbb{T}$ (α being a four-dimensional index - preferably an abstract index) gives us a complex-linear mapping from the space $\mathcal{F}_n$ of massless free fields of spin $\frac{1}{2}n$ to the corresponding space $\mathcal{F}_{n-1}$ for spin $\frac{1}{2}(n-1)$. Applying this in two steps we find (since the ordering of the steps is immaterial) that we have a (complex-linear) mapping from $\mathcal{F}_n$ to $\mathcal{F}_{n-2}$ given (linearly) by each element $S^{\alpha\beta}(= S^{\beta\alpha})$ of $\mathbb{T}\,\theta\,\mathbb{T}$ (where θ denotes symmetric tensor product). The space $\mathbb{T}\ \theta\ \mathbb{T}$ is ten-complex-dimensional (since $\frac{1}{2}\,(4 \times 5) = 10$). Now for a given 2-surface $\mathcal{S}$, the charge integral gives a linear mapping from $\mathcal{F}_2$ to $\mathbb{C}$, where we take $\mathcal{F}_n$ to refer to fields which need be defined only in the neighbourhood of $\mathcal{S}$. Here, the charge integral is being applied to the anti-self-dual field

$$^{-}F_{ab} = \phi_{AB}\epsilon_{A'B'} \qquad (4)$$

instead of to $^{*}F_{ab}$ and the result is

$$-\tfrac{1}{2}\,(\mu+ie) \qquad (5)$$

where μ and e are the magnetic and electric charges surrounded by $\mathcal{S}$, respectively. Composing the "twistor" map from $\mathcal{F}_4$ to $\mathcal{F}_2$

with the "charge integral" map from $\mathcal{F}_2$ to $\mathbb{C}$, we see that each element of $\mathbb{T} \theta \mathbb{T}$ gives a linear mapping from $\mathcal{F}_4$ to $\mathbb{C}$ providing, in the general case, ten independent complex "charge" integrals (compare Sachs and Bergmann[9]). Thus, the collection of these "charges", given $\mathcal{Y}$ and a specific element of $\mathcal{F}_4$ (defined in a neighbourhood of $\mathcal{Y}$), constitutes an element $A_{\alpha\beta}(= A_{\beta\alpha})$ of $\mathbb{T}^* \theta \mathbb{T}^*$ called the angular momentum twistor or moment twistor. For general spin $\frac{1}{2}n$ ($\geqslant 1$) we would have a complex charge integral for each element of $\theta^{n-2}\, \mathbb{T}$, giving, in all, $\frac{1}{2}(n-2)(n-1)$ independent complex charges which together constitute an element $Q_{\alpha_1\beta_2\cdots\alpha_{n-2}}\ (= Q_{(\alpha_1\cdots\alpha_{n-2})})$ of $\theta^{n-2}\, \mathbb{T}^*$.

It should be remarked that the dual twistor space $\mathbb{T}^*$ is naturally identified with the complex conjugate twistor space $\overline{\mathbb{T}}$. The elements of $\mathbb{T}^*$ are solutions of the conjugate twistor equation

$$\nabla_A^{(A'} \mu^{B')} = 0 \ . \tag{6}$$

It is easily verified from (1) and (6) that the scalar quantity

$$\tfrac{1}{2} i(\mu^{A'} \nabla_{AA'} \omega^A - \omega^A \nabla_{AA'} \mu^{A'}) \tag{7}$$

is constant throughout $\mathcal{M}$. This constant is the scalar product of the element of $\mathbb{T}$ with the element of $\overline{\mathbb{T}}$ and serves to identify $\overline{\mathbb{T}}$ with $\mathbb{T}^*$.

If we denote the spinor field ω^A, which satisfies (1), by Z^α (α being a four-dimensional abstract index) and the spinor field $\mu^{A'}$ satisfying (6) by W_α, then we denote the complex number (7) by

$$Z^\alpha W_\alpha \tag{8}$$

this being the twistor scalar product. Note that $\bar{\omega}^{A'}$ would be a particular solution of (6), which we denote by $\bar{Z}_\alpha$, and $\bar{\mu}^A$ a particular solution of (1), which we denote by $\bar{W}^\alpha$. The operation of twistor complex conjugation reverses the up/down positioning of twistor indices. The quantity $Z^\alpha \bar{Z}_\alpha$ (or $\bar{W}^\alpha W_\alpha$) is referred to as the twistor "norm". It is not positive definite, but has signature ++--.

For an electromagnetic field there is a reality condition on

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

the complex charge $\mu + ie$ —asserting the vanishing of the magnetic charge μ —whenever the field is obtainable in the standard way from a (single) potential. Analogously, we expect some kind of corresponding Hermiticity condition to apply to $A_{\alpha\beta}$ whenever the linearized curvature K_{abcd} is obtained in the standard way from a potential, namely from a linearized metric h_{ab}. This is indeed the case, but unlike the electromagnetic condition the (linearized) gravitational one is not conformally invariant. For its expression, one needs the <u>infinity twistor</u> $I^{\alpha\beta}(= -I^{\beta\alpha})$, an element of $\mathbb{T} \wedge \mathbb{T}$ which is <u>simple</u>

$$I^{\alpha[\beta} I^{\gamma\delta]} = 0 \qquad (9)$$

and whose dual

$$I_{\alpha\beta} = \tfrac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} I^{\gamma\delta} \qquad (10)$$

is equal to its complex conjugate

$$\bar{I}_{\alpha\beta} = I_{\alpha\beta} \ . \qquad (11)$$

Then the required Hermiticity condition on $A_{\alpha\beta}$ is

$$A_{\alpha\beta} I^{\beta\gamma} = \bar{A}^{\gamma\beta} I_{\beta\alpha} \qquad (12)$$

(i.e. $A_{\alpha\beta} I^{\beta\gamma}$ is "Hermitian" in the twistor sense).

It is frequently helpful to refer to the <u>spinor parts</u> of a twistor. These are various special spinor fields, connected by certain differential relations, which are associated with a twistor and which, at any one point of $\mathbb{M}$, determine the twistor completely. For a twistor of valence k there are 2^k spinor parts each of which is a spinor field of valance k but with varying numbers of unprimed and primed indices. The spinor parts of Z^α, defined by ω^A above, are ω^A itself and the constant field $\pi_{A'}$ given by

$$\nabla_{AA'} \omega^B = -i \varepsilon_A{}^B \pi_{A'} \ , \qquad (13)$$

the form of this expression being a consequence of (1). We write

$$Z^{\alpha} = (\omega^{A}, \pi_{A'}) \ . \tag{14}$$

Correspondingly W_{α} has spinor parts $\mu^{A'}$ and λ_{A}, with

$$\nabla_{AA'}\mu^{B'} = i\varepsilon_{A'}{}^{B'}\lambda_{A} \tag{15}$$

(see (6)) and we write

$$W_{\alpha} = (\lambda_{A}, \mu^{A'}) \ . \tag{16}$$

The scalar product (7) now becomes

$$Z^{\alpha}W_{\alpha} = \omega^{A}\lambda_{A} + \pi_{A'}\mu^{A'} \ . \tag{17}$$

The spinor parts of $I_{\alpha\beta}$ and $I^{\alpha\beta}$ are given by

$$I_{\alpha\beta} = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon^{A'B'} \end{pmatrix} \ , \quad I^{\alpha\beta} = \begin{pmatrix} \varepsilon^{AB} & 0 \\ 0 & 0 \end{pmatrix} \tag{18}$$

and those of $A_{\alpha\beta}$ by

$$A_{\alpha\beta} = \begin{pmatrix} 0 & p_{A}{}^{B'} \\ p^{A'}{}_{B} & 2i\mu^{A'B'} \end{pmatrix} \ . \tag{19}$$

Here p_{a} denotes the 4-momentum surrounded by $\mathscr{S}$; and M^{ab}, given by

$$M^{AA'BB'} = \bar{\mu}^{AB}\varepsilon^{A'B'} + \varepsilon^{AB}\mu^{A'B'} \ , \tag{20}$$

denotes the 6-angular momentum surrounded by $\mathscr{S}$, taken about the field-point at which $\mu^{A'B'}$ is being examined. The particular form (19), with a zero in the upper left corner and with

$$p_{a} = \bar{p}_{a}, \quad \mu_{A'B'} = \mu_{B'A'} \ , \tag{21}$$

follows from (12) together with the symmetry

$$A_{\alpha\beta} = A_{\beta\alpha} \ . \tag{22}$$

The fact that p_{a} is a constant and that the $M^{\alpha\beta}$ defined by (20) satisfies

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

$$\nabla_a M^{bc} = 2\, g_a{}^{[c}\, p^{b]} \tag{23}$$

(which is a general property of the dependence of angular momentum on position) follow from the generalizations of (15) that hold for 2-valent twistors. We can obtain the rest-mass m, surrounded by $\mathcal{S}$, directly from the formula

$$m^2 = -\tfrac{1}{2} A_{\alpha\beta} \bar{A}^{\alpha\beta} \tag{24}$$

which follows from (19) and the rules for twistor complex conjugation.

If ξ^a is a (real) Killing vector for $\mathcal{M}$, so that

$$\nabla_{(a}\, \xi_{b)} = 0, \tag{25}$$

then the dependence of ξ^a on position is such that it has the form

$$\begin{aligned} \xi^{AA'} &= \tfrac{1}{3}(i\nabla_B^{A'} \sigma^{AB} - i\nabla_{B'}^{A} \bar{\sigma}^{A'B'}) \\ &= -\tfrac{1}{6}\, e^{AA'bcd} \nabla_b W_{cd} = -\tfrac{1}{3}\, \nabla_b {}^{*}W^{ab} \end{aligned} \tag{26}$$

where e^{abcd} is the alternating tensor and

$$W^{ab} = -W^{ba}, \quad \sigma^{AB} = \sigma^{BA}, \tag{27}$$

$$W^{AA'BB'} = \sigma^{AB} \varepsilon^{A'B'} + \varepsilon^{AB} \bar{\sigma}^{A'B'}, \tag{28}$$

and where

$$\nabla_{A'}^{(A}\, \sigma^{BC)} = 0 \tag{29}$$

i.e.

$$\nabla^{(a}\, W^{b)c} - \nabla^{(a} W^{c)b} + g^{a[b} \nabla_d W^{c]d} = 0 \tag{30}$$

(see Penrose and Rindler[5]).

Equation (29) is the valence-2 version of the "twistor equation" (1) for a symmetric twistor $S^{\alpha\beta} (= S^{\beta\alpha}) \in \mathbb{T} \odot \mathbb{T}$, and we have, for its spinor parts

$$S^{\alpha\beta} = \begin{bmatrix} \sigma^{AB} & \rho^{A}{}_{B'} \\ \rho_{A'}{}^{B} & \tau_{A'B'} \end{bmatrix} \tag{31}$$

where

$$\nabla_{CC'}\sigma^{AB} = -2i\varepsilon_C{}^{(A}\rho_{C'}^{B)} \tag{32}$$

$$\nabla_{CC'}\rho_{B'}^{A} = -i\,\varepsilon_C{}^{A}\tau_{C'B'} \tag{33}$$

$$\nabla_{CC'}\,\tau_{A'B'} = 0 \tag{34}$$

and where $\tau_{A'B'}$ is symmetric. From (26) and (32) we get

$$\xi^a = -2\mathrm{Re}\{\rho^a\} \tag{35}$$

and ρ^a is, in fact, a complex Killing vector.

The Killing vector ξ^a picks out a particular component of energy-momentum/angular momentum, namely

$$\mathrm{Re}\{S^{\alpha\beta}A_{\alpha\beta}\} = \rho_{B'}^{A}p_A^{B'} + \bar{\rho}_{B'}^{A}p_A^{B'} + i\tau_{A'B'}\mu^{A'B'} - i\bar{\tau}_{AB}\bar{\mu}^{AB}$$

$$= \xi^a p_a + \tfrac{1}{2}(\nabla_a\xi_b)M^{ab} \ . \tag{36}$$

For each choice of $S^{\alpha\beta} \in \mathbb{T} \otimes \mathbb{T}$, this gives us the required particular component of $A_{\alpha\beta}$. Because of (12), there is a gauge freedom in $S^{\alpha\beta}$ in the sense that

$$S^{\alpha\beta} + 2iG_{\gamma}^{(\alpha}I^{\beta)\gamma} \ , \tag{37}$$

where G_β^α is Hermitian in the twistor sense, yields the same component in (36) as does $S^{\alpha\beta}$.

IN CURVED SPACE-TIME

In a general curved space-time $\mathcal{M}$, the equation (1) does not have any non-trivial solutions. However, the idea is to apply the spin-lowering procedure only at the surface $\mathcal{S}$ itself[1]. This will give us a plausible definition of the energy-momentum/angular momentum surrounded by $\mathcal{S}$, but it will not immediately indicate how we may obtain any kind of conservation law. This is all that we need at this stage. In any case, we do not want a conservation law of the strong kind that is obtained in Maxwell theory. In that theory, the charge would remain constant whenever $\mathcal{S}$ is moved continuously through any region of space-time where the charge-

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

current vector vanishes - so the field does not act as its own source. In the non-linear Einstein theory we anticipate that the field does to some extent act as its own source. (Recall the opening remarks concerning gravitational wave energy and potential energy.) If we had a full set of solutions of (1), then such a strong conservation law would be a consequence. It should be noted, on the other hand, that a weak form of conservation law is automatically incorporated in the fact that the definition refers only to the 2-surface $\mathscr{S}$ and the field-quantities defined at $\mathscr{S}$, since the "charges" obtained are independent of how one might choose to span $\mathscr{S}$ by a 3-surface.

The restriction of (1) to $\mathscr{S}$ can be written

$$\eth'\omega^0 = \sigma'\omega^1, \quad \eth\omega^1 = \sigma\omega^0, \tag{38}$$

where the compacted spin-coefficient formalism[10),8)] is being adopted, the spin-frame (o^A, ι^A) being chosen so that the flagpole of o^A is the outgoing null normal at each point of $\mathscr{S}$, and that of ι^A is the ingoing null normal. With $\mathscr{S}$ a spacelike 2-surface with S^2 topology, the differential equations (38) always have at least four independent solutions; in the general case they have exactly four[5)]. Let us assume the number is indeed four. Then we refer to the complex four-dimensional space $\mathbb{T}(\mathscr{S})$ of solutions of (38) as the 2-surface twistor space defined by $\mathscr{S}$.

According to our procedure outlined above, the components of total energy-momentum/angular momentum surrounded by $\mathscr{S}$ are obtained by "snuffing off" two of the indices of R_{abcd}, using quantities W^{ab}, spanned (over $\mathbb{R}$) by those which take the form

$$W^{AA'BB'} = \omega^A\omega^B\varepsilon^{A'B'} + \varepsilon^{AB}\bar{\omega}^{A'}\bar{\omega}^{B'} \quad . \tag{39}$$

We then define the total moment twistor $A_{\alpha\beta}$ for the quantities surrounded by $\mathscr{S}$ by

$$A_{\alpha\beta}Z^\alpha Z^\beta = -\frac{1}{16\pi G}\oint_{\mathscr{S}} \eta\, R_{AA'BB'cd}\,\omega^A\omega^B\varepsilon^{A'B'}dx^c \wedge dx^d \tag{40}$$

where, for the present, we take $\eta = 1$. (The inclusion of an η different from unity has not been motivated by the discussion so far,

but the modification referred to in the introduction requires this factor, as we shall see later.) Here, G is Newton's gravitational constant, and the 2-surface twistor Z^α is the element of $\mathbb{T}(\mathcal{J})$ defined by the chosen solution ω^A of (38). We take $A_{\alpha\beta}$ symmetric, i.e. $A_{\alpha\beta} \in \mathbb{T}^*(\mathcal{J}) \otimes \mathbb{T}^*(\mathcal{J})$.

The energy-momentum components of expression (40) (with $\eta = 1$) can be shown[1),5)] to yield precisely the Bondi-Sachs 4-momentum[11),12)] when $\mathcal{S}$ is taken to be a cut of $\mathcal{J}^+$. The angular momentum components of $A_{\alpha\beta}$, however, provide a somewhat different (and more satisfactory) set of quantities than had been defined previously. For $\mathcal{S}$ at spacelike infinity i^0, the definition (40) has been shown by Shaw[13)] to yield both the standard Arnowitt-Deser-Misner[14)] definition of energy-momentum and the Ashtekhar-Hansen[15)] definition of angular momentum.

More recently, R.M. Kelly[16)] has examined the behaviour of (40) in space-times which are asymptotic to anti-de Sitter space in the appropriate sense. Since the relevant field equations for such space-times are Einstein's equations with a cosmological constant λ, we must replace the Riemann tensor in (40) according to the substitution

$$R_{ab}{}^{cd} \to R_{ab}{}^{cd} - \frac{2}{3}\lambda\, g_{[a}{}^{c}\, g_{b]}{}^{d} \,. \qquad (41)$$

With this alteration, the Abbott-Deser[17)] definitions of "energy-momentum/angular momentum" are obtained.

SOME EXAMPLES FOR FINITE SURFACES

Tod[3)] has looked at a number of different examples for finite surfaces $\mathcal{S}$. His main procedure involves finding an embedding of $\mathcal{S}$ in a conformally flat space-time $\mathcal{M}$ in such a way that the intrinsic metric of $\mathcal{S}$ and the extrinsic curvatures of $\mathcal{S}$ are not altered from what they had been in $\mathcal{M}$. The 2-surfaces for which such an embedding is possible are called uncontorted, while those for which such an embedding is not possible are colled contorted.

Tod finds that if $\mathcal{S}$ is drawn arbitrarily on a hypersurface of revolution

$$t = f(r) \tag{42}$$

in the Schwarzschild space-time, where t and r are the usual Schwarzschild coordinates and f is an arbitrarily chosen smooth function, then $\mathcal{S}$ is uncontorted. Moreover, he finds that for such 2-surfaces the mass m, as calculated by the formula (24), gives precisely the Schwarzschild mass when $\mathcal{S}$ links the source (once) and gives zero whenever $\mathcal{S}$ does not link the source. A similar statement holds when $\mathcal{M}$ is the Reissner-Nordstrom space-time except that now the value of m is modificed by an appropriate-looking integral of electromagnetic field energy (and which does not suffer from the "factor-of-two" problems that plague the Komar integral expression[18]).

It is natural, in view of the above results, to conjecture that if $\mathcal{S}$ is drawn arbitrarily in the Schwarzschild space-time, then m should vanish whenever $\mathcal{S}$ does not link the source and should be equal to the Schwarzschild mass whenever it does link the source (once). However, in the general case $\mathcal{S}$ is contorted and this leads to various complications. In the uncontorted cases there is a natural identification between $\overline{\mathcal{T}}$ and $\mathcal{T}^*$ (which is conformally invariant) and this leads to an unambiguous definition of the expression $A_{\alpha\beta}\tilde{A}^{\alpha\beta}$ that is needed for the definition (24) of m^2. However, when $\mathcal{S}$ is contorted, it is much less clear how these procedures are to be carried out. It turns out, also, that in such cases we need to worry about the η-factor that was introduced without motivation in the expression (40). This matter will be returned to in the final section.

Another example considered by Tod concerns space-times in which there is a moment-of-time-symmetry hypersurface $\mathcal{H}$ which is intrinsically conformally flat. We assume that most of $\mathcal{H}$ lies in the vacuum region but that there are various localized regions $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$,... of matter on $\mathcal{H}$. (Alternatively, these matter regions can be replaced by "wormholes" or "funnels" connecting the hypersurface to another copy. This makes little difference to the discussion.) Tod finds that, for 2-surfaces $\mathcal{S}$ lying on $\mathcal{H}$, such surfaces are always uncontorted, and the value of m depends only

on the linking properties of $\mathcal{S}$ with the source regions. For example, m = 0 if no source is linked by $\mathcal{S}$. Let m_i be the value of m when just the i^{th} source region $\mathcal{R}_i$ is linked; let m_{ij} be the value when just the i^{th} and j^{th} source regions $\mathcal{R}_i$ and $\mathcal{R}_j$ are together linked. Tod finds the inequality

$$m_{ij} < m_i + m_j \ , \tag{43}$$

which is reasonable physically, where we recall that there will be a negative gravitational potential energy contribution to the total mass of a pair of sources, so this should indeed be less than the sum of the two constituent masses. It turns out, moreover, that in the limit of large separations the difference between m_{ij} and $m_i + m_j$ indeed agrees with the Newtonian value for the potential energy.

This result may be compared with a recent study, by B.P. Jeffryes, of the Newtonian limit for systems of many gravitationally interacting sources. Again agreement is obtained between expression (40) and the Newtonian potential energy contributions.

Another example considered by Tod is that of a static black hole surrounded by some framework which distorts the hole out of spherical shape. For $\mathcal{S}$ taken at the horizon, Tod finds the inequality

$$m^2 \geqslant \frac{A}{16\pi G^2} \tag{44}$$

(A being the horizon's surface area), the minimum mass, for given surface area, being achieved in the spherically symmetrical (Schwarzschild) case. This again is physically reasonable, since there will be an energy of distortion contributing to the black hole's mass when it is not spherical, and the surface area can be kept constant under slow (adiabatic) changes of shape. Again, $\mathcal{S}$ is uncontorted in these cases.

Surfaces in Robertson-Walker universes have also been examined by Tod[3]. Since the space-times are all conformally flat, any $\mathcal{S}$ within them must be uncontorted. For the closed models, with a spherically symmetrical $\mathcal{S}$ the value of m reaches a maximum

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

when $\mathcal{S}$ is "equatorial" and it reduces to zero when $\mathcal{S}$ constracts to a point on either side. For models with negative spatial curvature, there is an exponential increase of the total mass when $\mathcal{S}$ becomes very large.

These results seem perfectly reasonable physically, but a similar behaviour is also obtained by Tod (and also independently by Reula and Jeffryes) for the case of a 2-surface with a long cylindrically symmetric portion and squared-off-ends, in a cylindrically symmetric universe. There is either a sinusoidal or exponential behaviour for m, in terms of the length of the cylindrically symmetric portion of $\mathcal{S}$. At first sight the result seems physically strange, but there is a similarity with the cosmological results just cited, and on reflection the results are not perhaps unreasonable. Work is in progress.

SMALL SURFACES; THE CORRECTION FACTOR

Another calculation[7], first carried out by Tod, and then re-examined in detail by Kelly, refers to the situation when $\mathcal{S}$ is taken to be a small sphere of cross-section on a light cone $\mathcal{C}$ of an arbitrary point P in an arbitrary space-time $\mathcal{M}$. A timelike vector t^a is chosen at P; and an affine parameter r is chosen on each generator of $\mathcal{C}$, where r = 0 at P and where the gradient of r at P is normalized against t^a. The 2-surfaces $\mathcal{S}(r)$ are defined by r = const. (r small), and the quantities $A_{\alpha\beta} = A_{\alpha\beta}(r)$ and m = m(r) are then calculated for each $\mathcal{S}(r)$. Tod first found that for a general (non-vacuum) $\mathcal{M}$, the expected answer (in view of the linearized theory correspondence) is obtained for $A_{\alpha\beta}(r)$, at order r^3, namely that resulting from the volume contribution of the energy-momentum tensor T_{ab} within $\mathcal{S}$. When $\mathcal{M}$ is vacuum at P, a contribution to $A_{\alpha\beta}(r)$ first shows up at order r^5. The calculation had been modelled, to some extent on an earlier one by Horowitz and Schmidt[19] who had adapted an earlier suggestion due to Hawking[20] for a measure of the mass surrounded by $\mathcal{S}$. Horowitz and Schmidt also found a first contribution at order r^5 in vacuum, their expression yielding a positive-definite answer for the mass which arises from the Bel-Robinson tensor[8] at P. Tod's calculation

yields a somewhat similar expression but which is not positive definite. Indeed, the squared mass $(M(r))^2$ turns out sometimes to be negative! This happens, in particular, when $\mathcal{M}$ is Schwarzschild's space-time and t^a is taken not to lie in a (t,r)-plane of constant angular coordinates θ,ϕ. This seems highly unsatisfactory; moreover it contradicts the appealing conjecture, referred to in the previous section, that m should vanish for 2-surfaces $\mathcal{S}$ in the Schwarzschild space-time which do not link the source.

The 2-surfaces in question, for which this difficult arises, are all contorted. One is led, therefore, to consider the possibility of a modification to my original expression (40), in which η had been taken to be unity, this modification only coming into play for contorted 2-surfaces $\mathcal{S}$.

Independently of the Tod small spheres calculation I had been pursuing another line of thought aimed at proving the above-mentioned conjecture. My conclusion had been that if a certain property that I refer to as Property K always holds for $\mathcal{S}$ in Schwarzschild's space-time, then a form of that conjecture does hold true. Property K asserts that the Killing spinor[5] κ^{AB} which exists in that space-time, this being a solution of

$$\nabla_{A'}^{(A}\,\kappa^{BC)} = 0 \qquad (45)$$

when restricted to $\mathcal{S}$, is an element of $\mathbb{T}(\mathcal{S}) \otimes \mathbb{T}(\mathcal{S})$. This is a tempting speculation (partly because of the formal identity between (45) and (29)). Whenever Property K is true, it is possible to show, using a twistor-motivated contour integral expression, that for any $\mathcal{S}$ not surrounding the source in the Schwarzschild space-time $A_{\alpha\beta}$ must vanish, but where the η in equation (40) is not taken to be unity but proportional to

$$\begin{pmatrix} \omega^0 & \omega^1 & \eth\omega^0 & \eth'\omega^1 \\ 1 & 1 & 1 & 1 \\ \omega^0 & \omega^1 & \eth\omega^0 & \eth'\omega^1 \\ 2 & 2 & 2 & 2 \\ \omega^0 & \omega^1 & \eth\omega^0 & \eth'\omega^1 \\ 3 & 3 & 3 & 3 \\ \omega^0 & \omega^1 & \eth\omega^0 & \eth'\omega^1 \\ 4 & 4 & 4 & 4 \end{pmatrix} \qquad (46)$$

 Gravitational field mass. *Gravitational collapse and relativity; Yamada Conference XIV.* Eds H. Sato and T. Nakamura, 43–59. 1986

where ω_1^A, ω_2^A, ω_3^A, ω_4^A, are four independent solutions of (38).

The use of an η proportional to (46) can also be motivated in other ways. Indeed, when Tod and Kelly incorporated this η into their calculation they found that m(r) actually vanishes at order r^5 in any vacuum space-time. Moreover, according to a more sophisticated calculational procedure due to Woodhouse[7], small ellipsoids can also be treated (the light cone $\mathcal{C}$ being now replaced by an arbitrary spacelike hypersurface) and again it is found that m(r) vanishes at order r^5 when the above choice of η is made. This is a quite remarkable and non-obvious fact. (Indeed, for many months, one of the striking cancellations which allows this to happen had been overlooked, owing to an omitted term in the complicated calculation.) The above remarks also apply to the $p_a(r)$ and $\mu^{AB}(r)$ parts of $A_{\alpha\beta}$, but the status of the "zero" part of $A_{\alpha\beta}$ (upper left corner) is uncertain, as of now.

When $\mathcal{S}$ is uncontorted - and also when $\mathcal{S}$ is a contorted cut[5] of $\mathcal{I}^+$ - it turns out that η is always constant over $\mathcal{S}$ and so can be taken to be unity. Thus, the various good results that have been obtained are unaffected by the η-modification. However, in the general case there is a problem because one does not know which multiple of (46) to take for η. One possibility is to demand that η be normalized so that the average of its logarithm over $\mathcal{S}$ be zero; another is to demand that its own average be unity. This merely compounds two other difficulties that occur with general contorted surfaces, namely the problems of finding a natural correspondence between $\overline{\mathbb{T}}(\mathcal{S})$ and $\mathbb{T}^*(\mathcal{S})$ and for finding an appropriate analogue of the infinity twistor $I_{\alpha\beta}$, so that a version of the Hermiticity condition (12), and also an analogue of (24), can be found. There is reason to believe that, even in the general uncontorted case, such an $I_{\alpha\beta}$ will not satisfy (9) but would be an infinity twistor appropriate to an (anti-)de Sitter space-time[5] rather than to $\mathbb{M}$. All this represents work in progress.

REFERENCES

1. Penrose, Quasi-local mass and angular momentum in general relativity, Proc. Roy. Soc. Lond. A381, 53-63 (1982)

2. Penrose, R. Integrals for general-relativistic sources: a development from Maxwell's electromagnetic theory, in Maxwell Symposium Volume, ed. M.S. Berger (North-Holland, Amsterdam 1984)

3. Tod, K.P. Some examples of Penrose's quasi-local mass construction, Proc. Roy. Soc. Lond. A388, 457-77 (1983)

4. Tod, K.P. Three-surface twistors and conformal embedding, Gen. Rel. Grav. 16, 435-43 (1984)

5. Penrose, R. and Rindler, W. Spinors and Space-Time; Vol. 2, Spinor and Twistor Methods in Space-Time Geometry (Cambridge University Press, Cambridge 1986)

6. Penrose, R. Energy and its Definition in General Relativity in Twelfth Texas Symposium on Relativistic Astrophysics (Eds. M. Livio and G. Shaviv, Ann. N.Y. Acad. Sci. Vol. 470, 136-146 1986)

7. Kelly, R.M., Tod, K.P. and Woodhouse, N.M.J. Quasi-Local Mass for Small Surfaces, Class. Quantum Grav. (to appear)

8. Penrose, R. and Rindler, W. Spinors and Space-Time; Vol. 1, Two-Spinor Calculus and Relativistic Fields (Cambridge University Press, Cambridge 1984)

9. Sachs, R.K. and Bergmann, P.G. Structure of particles in linearized gravitational theory, Phys. Rev. 112, 674-80 (1958)

10. Geroch, R.P., Held, A. and Penrose, R. A space-time calculus based on pairs of null directions. J. Math. Phys. 14, 874-81 (1973)

11. Bondi, H. Gravitational waves in general relativity, Nature London, 186, 535 (1960)

12. Sachs, R.K. Gravitational waves in general relativity VIII. Waves in asymptotically flat space-time, Proc. Roy. Soc. Lond. A270, 103-26 (1962)

13. Shaw, W.T. Twistor theory and the energy-momentum and angular momentum of the gravitational field at spatial infinity, Proc. Roy. Soc. Lond. A390, 191-215 (1983)

14. Arnowitt, R., Deser, S. and Misner, C.W. Coordinate invariance and energy expressions in general relativity, Phys. Rev. 122, 997-1006 (1961)

15. Ashtekar, A. and Hansen, R.O. A unified treatment of null and spatial infinity in general relativity. I. Universal structure, asymptotic symmetries, and conserved quantities at spatial infinity, J. Math. Phys, 19, 1542-66 (1978)

16. Kelly, R.M. D.Phil. thesis, Oxford 1986

17. Abbott, L.F. and Deser, S. Stability of gravity with a cosmological constant, Nucl. Phys. B195, 76-96 (1982)

18. Komar, A.B. Covariant conservation laws in general relativity, Phys. Rev. 113, 934-6 (1959)

19. Horowitz, G.T. and Schmidt, B.G. Note on Gravitational Energy Proc. Roy. Soc. Lond. A381 215-224 (1982)

20. Hawking, S.W. Gravitational radiation in an expanding universe, J. Math. Phys. 9, 598-604 (1968)

164
GRAVITY AND QUANTUM THEORY

(Based on Lecture Delivered in Bangalore for Bangalore Pacific Asia Conference, 1986). *Proceedings of the Second Asia-Pacific Physics Conference* (Bangalore). Ed. S. Chandrasekhar. World Scientific, Singapore 1987

468 Gravity and quantum theory. (Based on Lecture Delivered in Bangalore for Bangalore Pacific Asia Conference 1986). *Proceedings of the Second Asia-Pacific Physics Conference* (Bangalore). Ed. S. Chandrasekhar. 1987

GRAVITY AND QUANTUM THEORY

R. Penrose
Mathematical Institute
Oxford, U. K.

It is generally recognised that the problem of quantum gravity is one of the most important frontiers of our current physical understanding. In saying this, I should make clear that superstring theory, despite its many novel and fascinating features, remains a long way from a solution of this problem. What has been achieved in the direction of quantum gravity with superstrings, is only at the perturbative level, whereas there are good reasons to believe that a conventional perturbative approach, starting with flat space, is unlikely ever to succeed. The same comment applies to the earlier supergravity theory. Somewhat more in the true spirit of general relativity is the approach being carried out in Cambridge by Hawking and his school. Yet that approach, also, lies a good way from its avowed goals.

Difficulties with standard general relativity, when a perturbative approach to quantization has been adopted, have led many physicists to believe that, even at the classical level, standard general relativity needs to be replaced by some other theory more amenable to the conventional methods of quantization. Supergravity, generalized Kaluza-Klein theories and superstring theories are cases in point. However my own attitude is quite different from this. I believe that there will be as much change in the structure of quantum theory as in that of general relativity, when the true theory uniting the two is eventually found (see also ref. 1).

My reasons for believing this lie partly in problems that arise when one tries to apply the technique of quantization even in a fixed curved-space background.

Quantization is not really well defined procedure in the context of general coordinate freedom. The difficulty boils down to the fact that the group of canonical transformations of a classical theory (symplectic manifold) is not the same as the unitary group of a quantum theory. That is to say, roughly speaking, one needs to have some special choices of coordinates for the space-time. These are to be the coordinates whose Poisson brackets with corresponding momenta are to be replaced by the canonical commutators of the proposed quantum theory. In general, where there are no translational space-time symmetries to define these momenta in a meaningful way, no clear quantization procedure seems to be singled out. In a proper quantum gravity theory, moreover, the geometries themselves have to be involved in quantum linear superposition. The linear superposition of states which involve significantly differing geometries must result in something which in no way resembles a "space-time geometry" in the ordinary sense. What would be the quantization rules in a background which itself is not even a space-time geometry? I think it is clear that we do not at all know what the "rules of quantization" actually <u>are</u> when gravity becomes involved in an important way - so to "change" these rules is not really such an outrageous step as might be thought!

It must also be said that, despite its profound mathematical beauty and its marvellous agreement with all experimental tests to date, the quantum

theory is hardly satisfactory as a complete coherent theory of the world; for it is a curious concoction of, on the one hand, deterministic unitary evolution and, on the other, probabilistic state-vector reduction. The theory makes no clear statement as to when unitary evolution must be suspended and the second process is to take over. Yet, without this second process, there would be no agreement with experiment.

There were, of course, many different attitudes to this dilemma. My own point of view is that both unitary evolution and state-vector reduction are real; that is to say, they are good approximations to some objective physical processes which are actually taking place in the physical world. The correct theory of these physical processes would have to be some non-linear (at least) generalization of standard quantum theory, the rough idea being that when the states to be superposed become in some well-defined sense "too different" from one another, a non-linear instability takes over and one of the states wins out over the other, rather than their both coexisting in linear superposition.

What is the appropriate criterion for the states being "too different"? My belief is that the appropriate physical input from this comes from general relativity. I have given elsewhere[2] my reasons (an investigation of cosmology, black holes and the second law of thermodynamics) for believing that there is an asymmetric law governing space-time singularity structure. This law is suggested to take the form that past singularities have vanishing Weyl conformal curvature whereas future singularities are unconstrained. The law would presumably be a deduction from some yet-to-be-discovered quantum gravity theory. The time-asymmetric nature of this law implies that such quantum gravity theory would have to be a time-asymmetric theory.

Gravity and quantum theory. (Based on Lecture Delivered in Bangalore for Bangalore Pacific Asia Conference 1986). *Proceedings of the Second Asia-Pacific Physics Conference* (Bangalore). Ed. S. Chandrasekhar. 1987

Phase-space considerations concerning "gedanken experiments" with isolated containers with perfectly reflecting walls suggest that the "other side of the coin" from the above Weyl curvature law is a gravitational theory of state-vector reduction[3,4,5]. The idea is that when a quantum state involves a superposition of two significantly differing space-time geometries, linearity fails, and perhaps as a result of some instability one or the other geometry is singled out. In this, "significantly different" means different, to the tune of at least one (longitudinal) graviton. This is only a very rough criterion, at the present state of understanding.

To illustrate what I have in mind, consider an atom which decays radioactively. A spherical wave emerges, but at some distance there is a detector, say a cloud chamber. The state of the emitted particle gets coupled to a number of different configurations (tracks) of condensing droplets. When the droplets reach a certain size, according to this scheme, one of these tracks gets singled out. To estimate when this should be, consider the simplified case of a single droplet. Suppose, as a crude idealization, the droplet is a sphere of radius r, which is condensed from a region of vapour of radius R, leaving vacuum in between. There is Weyl curvature in this vacuum (Schwarzschild solution) and this is the region which contributes to the (longitudinal) graviton number. An estimate of this number is

$$\frac{768\,\pi^3\,G}{\hbar c^3}\,.\,m^2\,\log\left(\frac{R}{r}\right)$$

(the coefficient $768\pi^3$ having been supplied by A.Ashtekar) where m is the mass of the droplet. If the logarithm is of order unity then $m \approx 10^{-7}$g for the one-graviton criterion, perhaps not an unreasonable figure.

Clearly much more work needs to be done if these ideas are plausible dynamical theory of state-vector reduction.

REFERENCES

1 Károlyházy,F. (1966) Nuovo Cimento **A42**, 390.

2 Penrose,R. (1979) in General Relativity: an Einstein Centenary, eds. S.W.Hawking and W.Israel (Cambridge University Press, Cambridge)

3 Penrose,R. (1981) in Quantum Gravity 2, eds. D.W.Sciama, R.Penrose and C.J.Isham (Oxford University Press, Oxford)

4 Penrose,R. (1986) in Quantum Concepts in Space and Time, eds. R.Penrose and C.J.Isham (Oxford University Press, Oxford)

5 Penrose, R. (1987) Newton, Quantum Theory and Reality - 200th Anniversary of Principia Volume, eds. S.W.Hawking and W.Israel. (Cambridge University Press, Cambridge).

165
TWISTORS AND THE TIME-IRREVERSABILITY OF STATE-VECTOR REDUCTION

Twistor Newsletter **22** (1986). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 Twistors and the time-irreversability of state-vector reduction. *Twistor Newsletter* **22** (1986). Also published in *Further Advances in Twistor Theory,* Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.29 Twistors and the Time-Irreversibility of State-Vector Reduction *by R. Penrose* (TN 22, September 1986)

Twistor theory—insofar as it is a physical theory at all—is really rather conservative. As it stands at present, it meekly accepts the rules of quantum theory without any serious attempt at significant modification. According to the scheme of twistor diagram theory, one computes probability amplitudes, by some (admittedly unconventional) "twistorial" procedures, yet the intention is merely that those amplitudes should ultimately be used to calculate quantum-mechanical probabilities in the standard way. Even where twistor theory is at its most radical in this kind of respect, namely, in the philosophy underlying the "nonlinear graviton construction", the expressed intention is to use those "graviton states"—via some (presumably nonlinear) "scalar product" expression—to calculate complex amplitudes whose squared moduli would again be used to provide quantum mechanical probabilities in the conventional way.

Of course, there *is* a good deal of the mathematics of standard quantum theory underlying twistor theory at the foundational level—and rightly so! The very complex numbers which lie at the root of the quantum-mechanical formalism (these same probability amplitudes!) are to be thought of, also, as the complex numbers out of which twistors are themselves to be constructed. Moreover, the quantization procedure $\overline{Z}_\alpha \to \partial/\partial Z^\alpha$ has a considerable role to play in twistor theory, as have the corresponding connections with geometric quantization, symplectic structure, Hamiltonians, etc. Yet if twistor theory purports (as evidently it does) to generate new insights into the way that quantum theory and space-time theory are to combine into some unified whole, then it should take into account what I consider to be the one known "observational" fact of quantum gravity theory, namely, that it should (I claim) be *time-asymmetric*—owing to the grossly time-asymmetric structure of space-time singularities (presumably a quantum gravity effect)! The only place I know of where an ingredient with a *potential* time-asymmetry has been seriously introduced into twistor theory is with the "k" that A.P. Hodges found was needed for his development of twistor diagram theory (see §I.5.10 and Hodges 1985). In Penrose (1985) I made a tentative proposal according to which *state-vector reduction* might occur spontaneously and objectively as a (nonlinear) quantum-gravity effect. In Penrose (1985) I argued strongly that state-vector reduction must be grossly time-asymmetric. State-vector reduction is, after all, the place where the (holomorphic) probability amplitudes get converted (non-holomorphically) into actual probabilities. Accordingly, the way in which twistor theory ought, I claim, to be modified in order to incorporate a time-asymmetric ingredient is at the stage where probability amplitudes get converted into probabilities—and this ought to be a stage where "quantum gravitational" effects become important.

A shift in viewpoint forced itself upon me since the article Penrose (1981) (mentioned in an addendum to Penrose 1987). In Penrose (1981) I suggested that some measure of gravitational entropy was needed and that when this entropy reached a significant value (of order unity or a little more), state-vector reduction would take place. Partly as a result of criticisms by Robert Wald, I came to believe that this point of view is untenable, but that instead it is the (longitudinal) *graviton number* that is relevant. If two states under linear superposition subsequently couple into the gravitational field—to the extent that the difference between the fields of the two states reaches the level of one (longitudinal) graviton—then reduction to one or the other of them takes place

objectively.

This idea is developed more extensively in Penrose (1985) and in Penrose (1989). However, even this viewpoint is now superseded by that expressed in Penrose (1993), and in further work in Penrose (1994, 1996, 1998).

References

Penrose, R. (1979) Singularities and time-asymmetry, in *General Relativity—an Einstein Centenary Survey*, eds. S.W. Hawking & W. Israel, Cambridge University Press, Cambridge, p. 584.

Penrose, R. (1981) Time asymmetry and quantum gravity, in *Quantum Gravity* 2, eds. C.J. Isham, R. Penrose & D.W. Sciama, Oxford University Press, Oxford.

Hodges, A.P. (1985) Proc. Roy. Soc. Lond., **A397**, p. 341–74, 375–96.

Penrose, R. (1985) Gravity and state-vector reduction, in *Quantum Concepts in Space and Time*, eds. C.J. Isham & R. Penrose, Oxford University Press, Oxford.

Penrose, R. (1987) Newton, quantum theory, and reality, in *300 Years of Gravity*, eds. S. W. Hawking & W. Israel, Cambridge University Press, Cambridge.

Penrose, R. (1989) *The Emperor's New Mind*, Chapter 7, Oxford Univ. Press, Oxford.

Penrose, R. (1993) Gravity and quantum mechanics, in *General Relativity and Gravitation 1992 (Córdoba)*, Inst. Phys., Bristol, 179–189.

Penrose, R. (1994) *Shadows of the Mind*, Oxford University Press, Oxford.

Penrose, R. (1996) On gravity's role in quantum state reduction, Gen. Rel. Grav., **28**, 581–600.

Penrose, R. (1998) Quantum computation, entanglement and state reduction, Phil. Trans. Roy. Soc. Lond., **A356**, 1927–1939.

166
GRAVITY AND STATE VECTOR REDUCTION

Quantum concepts in space and time (1984), 129–46. Oxford University Press, 1986

478 Gravity and state vector reduction. *Quantum concepts in space and time* (1984), 129–46. 1986

9

Gravity and state vector reduction

Roger Penrose

At the Nuffield-supported Oxford conference, held in March 1984, on which the chapters in this book are based, many stimulating ideas and illuminating results were put forward. It would surely be foolish for me to attempt to summarize them. Nevertheless, I am certain that they have had a definite influence on my own particular viewpoint. I hope that this influence will not seem to be too subtle to be noticeable as actual changes in my expressed attitudes! These I shall be elaborating shortly.

But first, in order to put people into an appropriate frame of mind, I am providing a picture (Fig. 9.1) showing a concrete realization of the dynamic collapse of a wave function ψ. Provided that the state being described by ψ is sufficiently tiny or uncluttered, then ψ can float freely without any need for support. But a big hefty state needs all sorts of crutches to keep it up. Yet all is to no avail—it collapses anyway! (I was stimulated to show this by noticing that in Abhay Ashtekar's talk, a ψ was missing in one of his transparencies. It occurred to me that ψ must have collapsed!)

Despite this introduction, I should make clear that the rest of what I have to say actually *is* in deadly earnest.

9.1. Quantum theory and objective reality

I should begin by expressing my general attitude to present-day quantum theory, by which I mean standard non-relativistic quantum mechanics. The theory has, indeed, two powerful bodies of fact in its favour, and only one thing against it. First, in its favour are all the marvellous agreements that the theory has had with every experimental result to date. Second, and to me almost as important, it is a theory of astonishing and profound mathematical beauty. The one thing that can be said against it is that it makes absolutely no sense!

Its nonsensical aspect is, of course, the strange picture of 'reality' that it presents, where systems are described as evolving according to one or other of two apparently incompatible procedures: deterministic unitary

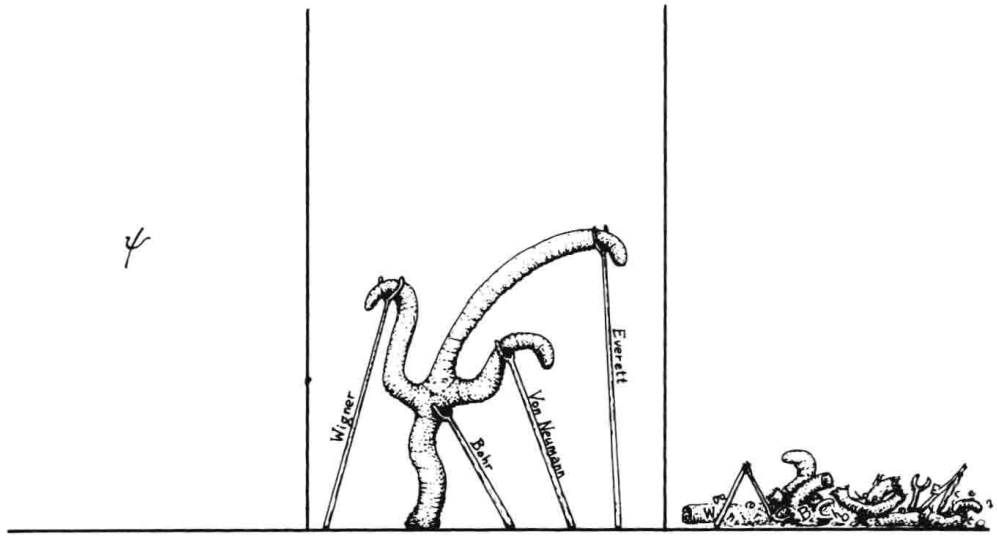


Fig. 9.1. Dynamical collapse of a wave function.

Schrödinger evolution (or, completely equivalently, Heisenberg evolution, but I shall, for convenience only, phrase my discussion entirely in the Schrödinger picture) and probabilistic state vector reduction. There is no clear-cut rule for deciding the stages at which unitary evolution must be suspended and replaced by the reduction procedure. This procedure has to have been invoked at least by the time an 'observation' is deemed to have been made by any conscious observer.† But state vector reduction may well be normally considered to be applicable a good while before this, at a stage where some 'irreversible' or 'macroscopic' change has been induced. Though there is no clear-cut rule telling us exactly when the reduction procedure is to be invoked, it is one of the powerful facts about the formalism of quantum mechanics that it makes *no difference* what stage the procedure is actually applied (cf. von Neumann (1932) and Chapter 6), provided that this is after any stage at which interference phenomena can plausibly be measured, and provided that it

† It is necessary to mention the many-worlds interpretation, however, (cf. Everett (1957) and Chapters 13 and 14), according to which reduction is supposed not to be mirroring anything which actually takes place physically at all, but which only appears to do so by some quirk of one's conscious perception. I have a number of serious reservations about this interpretation, but would not be appropriate for me to try to express these here (see Chapter 12 for a critique of the many-worlds view). Another viewpoint, in which, also, reduction is regarded as not taking place, assumes that the probabilistic behaviour normally described by the reduction procedure can be attributed to uncertainty in knowledge about the external environment, it being presumed that unitary Schrödinger evolution actually holds at all times. I do not regard this viewpoint as greatly plausible and I shall not attempt to give a discussion of it in this chapter.

is not beyond the stage of conscious perception. This powerful fact is indeed vital for the general applicability of the theory, and it obviates the need for any detailed description of a physical reduction procedure.

However, there is the unfortunate consequence, in my opinion, that people are easily led to take too subjective a view of the quantum mechanical picture of things. Since it seems to make so little difference when reduction is deemed to have taken place it is a commonly adopted view that in some way reduction is illusory and represents merely a shift in our knowledge about a physical system. (This leaves aside other possible attitudes, such as the many-worlds view according to which this illusion resides in the way a conscious observer would—allegedly—actually perceive just one world from among the linearly superposed infinitude of alternative possibilities.) The state vector itself is, accordingly, viewed as providing merely a 'state of knowledge' and not an objective description of (a part of) physical reality. The 'ensemble' viewpoint, according to which the state vector describes properties of ensembles of systems and not of a single physical system is similar, since again the state vector is taken as not describing an actual reality of the physical world. Such views illustrate a strong reluctance to take the state vector seriously—or sometimes, even, to take *reality* seriously!

However, it seems to me that so long as one accepts the mathematical procedures of quantum mechanics as valid for making predictions, one must also accept a certain physical objectivity for the state vector, at least as it applies to isolated systems. Let us envisage that the content of some isolating container C is initially prepared, at time $t=0$, in the (normalized) quantum state $|\psi_0\rangle$, and then is allowed to evolve, by unitary (Schrödinger) evolution to the state, at time t,

$$|\psi_t\rangle = U(t)\,|\psi_0\rangle \tag{9.1}$$

($U(t)$ being a unitary operator). The question I wish to raise is: 'does the state vector $|\psi_t\rangle$ provide an objective description of an actual *physically real* state within C?". It seems to me that the answer to this question, *provided* that one accepts the normal procedures of quantum mechanics, must be an unequivocal 'yes'; for at any time $t>0$ one has the option of performing an experiment which measures the (bounded) Hermitian operator

$$E_t = |\psi_t\rangle\,\langle\psi_t|\,. \tag{9.2}$$

The state vector $|\psi_t\rangle$ is (up to a phase) the *only* state with the property that with certainty it yields the result unity when the experiment is performed. The contents of C have to be 'ready' for the possibility that at any moment t, the experiment E_t *might* actually be performed—even when it is not going to be so measured (at least in any standard version of

the theory). So, it would seem to be clear that $|\psi_t\rangle$ (at least up to a phase multiplier) represents—at *any* time t (until some other measurement is made)—an *objective* property of the contents of C.

There is actually nothing conceptually problematical about this—as far as it goes. Why, then, is there such reluctance to accept the state vector as describing an actual physical reality? I am aware of at least three arguments. The first is the problem of 'Schrödinger's cat' (or some essential equivalent such as 'Wigner's friend', etc.) It seems hard to believe in any actual 'reality' at the level of a cat which requires that such states of death and life can co-exist. The situation is, of course, even more puzzling for Wigner's friend. Thus one may prefer to adopt the view that reality is no longer being described by the state vector, and the state vector now provides merely a calculational aid for computing probabilities of results to experiments. But what of our possible measurement E_t which *requires* that the cat be in a state neither dead nor alive? The standard argument would presumably be that, in this case E_t is a totally impossible measurement to perform in practice. But the formalism of quantum mechanics does not *in itself* have anything to say about what is practical and what is not. According to *strict quantum mechanics*, it seems to me, ψ_t—representing some complex linear combination of a live cat and a dead one—should still have to provide an objective description of a 'reality' contained within C.

My own view is that this is a clear indication that physical reality and the Schrödinger-evolved state vector have simply diverged from one another and, at the level of a cat, that state vector has indeed ceased to describe reality. Thus, on this view, the Schrödinger equation has become in fact grossly inaccurate in its description of reality. (There are, of course, many other views!) But the form of deviation from Schrödinger's equation that would be needed has a remarkable property, namely that the Schrödinger-evolved state vector remains, until the next 'observation' is made, an accurate tool for computing the probabilities of the possible outcomes of that 'observation'. For this reason, the 'mere computational tool' view of the state vector has its undoubted value.

A second argument against the objectivity of the state vector—or, more specifically, against the reality of state vector reduction—is that it seems not to be relativistically invariant. The difficulty is made manifest in experiments of the Einstein–Podolsky–Rosen type, and has been presented to us most forcefully as a fact of nature by the beautiful experiment of Aspect and his collaborators (cf. Chapter 1). Here one has a state consisting of a photon pair, the spins of which are observed at two events A, B with a *space-like* separation of some twelve metres. It would seem that the observation at each one of A and B should effect a state vector collapse, but the collapse which takes place *earlier*—say that at

A—provides the state which is to be observed in the *other* observation—say at B. However, the *temporal order* in which these collapses take place—i.e. whether A or B is considered to be the 'earlier'—depends upon the overall references frame and, therefore, according to relativity, is not an objective property. (In this I am making the assumption that each collapse takes place simultaneously in this 'overall reference frame'. If the collapses were to take place along, say, the future light cones of A and B then there would be an inconsistency with the correlations that Aspect and his co-workers actually observed, since such retarded collapse would be equivalent to using a *local* hidden variable theory.)

It does, however, make *no difference* to the correlations whether the collapse due to A or that due to B is deemed to have occurred 'first'. There is no inconsistency of the actual observations with relativity whichever choice is taken. There is, nevertheless, an inconsistency with the *spirit* of relativity *whichever* choice is taken if we wish to regard these collapsing state vectors as providing a picture of physical reality. One could, of course, simply ignore this fact and take the view that 'physically real' collapse takes place in some 'absolute' frame, determined perhaps cosmologically. This would be consistent with relativity, but it is a viewpoint that I should, myself, be most reluctant to adopt. It seems to me that the reduction process—which I believe to be, in some appropriate sense, *real*—must involve some very subtle and non-local space–time (or perhaps 'beyond' space–time†) ideas (cf. also Chapter 12).

Similar remarks may apply in relation to the third argument against state vector objectivity that I have in mind. This concerns the question of time-reversibility. The Schrödinger-evolution of the state vector which is normally envisaged (and which was argued for in the above discussion supporting state vector objectivity) requires that the evolution is such that the state is an eigenstate of the observation performed at the *beginning* of this evolution, not of the one at the end. This description is not time-symmetric; but precisely the same relative probabilities would be obtained if one took the 'perverse' view that it is the backwards-Schrödinger-evolved state vector that is relevant, this being the appropriate eigenstate of the operator describing the observation at the *future* end

† This kind of non-locality had, quite early on, supplied some of the motivations for the point of view of twistor theory (compare Penrose (1986)). It should be pointed out, however, that the violations of Bell's inequality that these experiments confirm, do not *in themselves* have anything to do with space–time or non-locality; it is just these particular *manifestations* of the violations of the inequality which do. But it is such manifestations that are most puzzling to our intuitions.

On this general topic, I might raise a further issue which occurred to me. In relation to Lee Smolin's talk, in which non-local hidden variables referring to *pairs* of particles were described, one might ask whether Bell-type inequalities for particle *triples* (or quadruples etc.) might exist for such theories, which could be violated in some quantum-mechanical experiment.

of the evolution, rather than the past (see Penrose (1979), p. 584). If we believe that time-symmetry ought to be a feature of an objective state vector reduction, then this might present us with a problem, over and above that presented in the preceding paragraph. However, my own view, as I shall present in the next section, will be seen to involve an essential time-asymmetry—and my preference will be that, to a good degree of approximation, physical reality is described by state vectors evolving in the normal future direction from some 'eigenstate' until some (spontaneous) reduction takes place.

9.2. Singularities and the Second Law

An essential feature of the physics involved in the reduction process must be the Second Law of Thermodynamics. If it is impossible *in principle* to measure, for example, the unitarily evolved state which represents Schrödinger's half dead and half alive cat, then it is because somehow the essential phase relations have become 'irretrievably lost'. One envisages that these phase relations have become spread over a large number of degrees of freedom; so, in some appropriate sense the entropy has 'gone up'. This irreversibility, however, has to be of some absolute and objective character if the phase relations are to be actually 'lost' in principle and not simply retrievable only with extraordinary difficulty. Thus, it would seem that we require some entropy concept with a greater physical objectivity than those that one normally encounters in physics.

In my view, an essential clue to this is to be found in the Bekenstein–Hawking *black hole entropy*, which is given by the remarkable formula

$$S_{\text{bh}} = \frac{kAc^3}{4\hbar G}, \tag{9.3}$$

where A is the surface area of the black hole's even horizon, and where c, G, k, and $\hbar$ are, respectively, the velocity of light, the gravitational constant, Boltzmann's constant, and Planck's constant divided by 2π. Note that here an entropy measure is set equal to a precise and seemingly *objective* quantitative feature of space–time geometry. I shall try to argue that the black hole entropy, as given by eqn (9.3), represents merely an extreme case of a *gravitational entropy*; and that it is this concept of gravitational entropy that has the needed objectivity and which supplies the physical ingredient involved in the triggering of state vector reduction. In order to motivate this claim, I shall need to outline a somewhat roundabout argument that I have given on earlier occasions concerning the origin of the Second Law, and its relation to the structure of space–time singularities. For further details I refer the reader to earlier publications (Penrose 1979, 1981).

The basic puzzle that the Second Law presents us with is, it seems to me, that at some time in the distant past, some very strong constraint must have been placed on the structure of the universe. This constraint resulted in a universe with a very low entropy, and the universe has been evolving to the more 'probable' situations of higher entropy every since—hence the Second Law. The time at which this constraint was imposed was presumably that of the Big Bang. We must ask what the nature of this constraint was, and try to understand it in relation to other aspects of physics.

One point that should first be made is that if one assumes deterministic evolution, then, in a certain sense, the state of the universe remains just as 'peculiar' at late times as it was at early times. But at late times this peculiarity takes the form of very precise *correlations* between particle motions. These correlated motions, when evolved backwards in time, produce the most manifest type of peculiarity, at early times, that is described as 'low entropy'. Thus, the form of peculiarity that our universe exhibits is, according to the Second Law: low entropy at early times, and precise correlations at late times. My point of view will always be to regard the presence of such correlations as secondary to (i.e. a consequence of) the low entropy, and not the other way about. Had it been the case in our universe that correlations of the opposite type which *reduce* entropy were *frequently* present (e.g. shattered glasses of water miraculously assembling themselves, etc., being commonplace occurrences) then I believe we should naturally have adopted a teleological type of terminology which asserts that the future low-entropy states are the *causes* of the earlier correlations (e.g. that the future existence of an assembled water glass provides the *reason* for the earlier precise particle motions in the pieces of glass and split water). Of course, *our* world is not like that, which is why this kind of teleological terminology seems unnatural. However it is a helpful form of words for the discussion of the Second Law, since it is not biased with regard to one direction of time or the other.

It is frequently argued that the Second Law arises *because* of the absence of precise entropy-reducing correlations in the initial state—and that what needs to be explained is the absence of such correlations. But my own view, as is implicit in the above, is that the presence or absence of such correlations is *secondary* to the entropy question, and is not the place to look for an explanation of the Second Law. In any case, it is the *low-entropy* (or manifest peculiarity) of the initial state which needs explanation. Given that, the absence of correlations which might reduce the entropy still further is no surprise. The vast majority of states compatible with the initial low entropy must indeed be uncorrelated in the sense that higher entropy states then evolve from them. To evolve to

even lower entropy, the initial state would have had to have been even more special than the requirement of initial low entropy already constrains it to be. That would have needed, in my way of looking at it, some further macroscopic constraint to have been imposed at some *later* time.

One apparent paradox is the fact that the Big Bang, though the universe-state of *lowest* entropy (by the Second Law), is treated as a state in which the matter is in thermal equilibrium, i.e. at *maximum* entropy! It should be made clear that the resolution of this paradox does *not* lie simply in the fact that the universe is expanding. (That is a commonly believed fallacy!) If one imagines the universe to be collapsing, one can still envisage that the Second Law holds. Galaxies fall in towards one another; they unite; the black holes at their cores unite; their stars fall together into more and more massive black holes, which all unite together into one final all-embracing singularity—the *Big Crunch*. Here the final singularity represents maximum entropy for the universe. Reversing time we obtain an expanding universe with a Big Bang whose entropy is a *maximum* and for which the opposite of the Second Law holds. Thus, the expansion of the universe can in no way explain the rise in entropy and is thus not at all the resolution† of the aforementioned paradox.

The actual resolution of the paradox lies in the fact that *gravitational degrees of freedom* (i.e. Weyl conformal curvature) are not taken into account in the 'thermal equilibrium' that seems to have been closely approximated in the very early stages of our actual universe. In the final stages of a generic 'Big Crunch', the Weyl curvature would diverge to infinity (as it does at the singularity even of the symmetrical Schwarzschild black hole), whereas the closely (Friedmann–) Robertson–Walker form that is assumed for the Big Bang has a Weyl curvature which is set equal to zero. Thus, I propose that the constraint on the Big Bang which sets the Second Law on its course is that initially the Weyl curvature must be set to zero.

This hypothesis has the implication that not only has the early universe its initial low‡ entropy, as is required for the Second Law, but the

† There seems to be a current misconception that the inflationary models somehow provide an explanation for the evident specialness of the initial state. But this is just another version of the same fallacy. These models all rely on time-symmetric physics, and the above argument still holds. There are vastly more modes of disappearance for a collapsing universe into a final singularity then those modes which resemble the time-reverse of an 'inflationary scenario'. Inflation does not explain why those vastly more probable modes of *reappearance* do not occur.

‡ Again there appears to be a common fallacy. The entropy per baryon figure of 10^8k–10^9k that is observed in the microwave background is often regarded as a 'high' value, whereas for a universe model with 10^{80} baryons, say, the entropy per baryon figure rises to $10^{43}k$ in the Big Crunch. Of course, an explanation for the observed figure is still needed.

uniformity and isotropy that is observed on a large scale is also a rough consequence. With gravitation it is the uniform states which have low entropy and the clumpy ones, high. Initially the gravitational contribution to the entropy is set to zero, but the entropy of the matter is at its maximum. Owing to the extreme weakness of the gravitational coupling, it is a long while before this entropy imbalance can take effect. But gradually the matter begins to clump as the gravitational entropy contribution rises. Galactic masses are formed, and matter collects into stars. The hot spot in our sky which is our Sun, and upon which all life on this planet depends for a needed powerful entropy sink, owes its existence to the clumping effect of gravity. The clumping together of masses rather greater than our Sun leads to black holes, these achieving the maximum entropy possible for a given mass (in a limited volume). For mass m, the maximum surface area for a black hole's event horizon is the Schwarzschild value

$$A = 16\pi G^2 c^{-4} m^2, \tag{9.4}$$

which, with the formula (9.3), yields the entropy

$$4\pi\hbar^{-1} k G c^{-1} m^2. \tag{9.5}$$

Note that it is entropically favourable for collections of small black holes to collect together into larger ones. This is consistent with the Second Law and our picture of the final collapse to a Big Crunch.

This is perfectly satisfactory as far as it goes. But it seems to me to be unsatisfactory simply to postulate for the Big Bang that its Weyl curvature should vanish, as a once-only condition unrelated to the other constraints given to us as physical laws. Space–time singularities are places where 'the known laws of physics break down'. But I do not believe that physics itself breaks down at a space–time singularity. It is just that the laws that govern their structure are presently unknown to us. The very uniformity of the structure of the Big Bang singularity is in itself an indication of the existence of a powerfully effective law. However this law cannot apply to the singularities of collapse, since that would lead to a violation of the Second Law and to the existence of peculiar teleological effects preventing the formation of local black holes. It is perhaps significant that (assuming Cosmic Censorship holds true; cf. Penrose 1978, 1979) space–time singularities can be separated into two disjoint classes: *past* singularities (singular TIFs) from which matter and signals can escape—e.g. the Big Bang or the singularities of white holes—and *future* singularities (singular TIPs) into which matter and signals can fall—e.g. the Big Crunch or the singularities of black holes. Thus, rather than applying the hypothesis of vanishing Weyl curvature specifically at

the occasion of the Big Bang and nowhere else, we can hypothesize that it applies in a more uniform universal way:

HYPOTHESIS: *The Weyl curvature vanishes at all past singularities, as the singularity is approached from future directions.*

This has the advantage that white holes, with their unpleasant anti-thermodynamic behaviour, are excluded. (The qualification about approaching from future directions is to attempt to cover possible violations of Cosmic Censorship that could occur at the final explosion in the Hawking evaporation of a black hole; cf. Penrose (1981).)

This hypothesis is time-asymmetric, as indeed could have been anticipated, since it yields the time-asymmetric Second Law. I shall indicate, moreover, that this time-asymmetry cannot be regarded as separate from all the other procedures of physics, so one must also anticipate some role for time-asymmetry elsewhere.

Envisage the following *gedanken* situation (introduced by Hawking 1976). Consider a large isolating container—with perfectly reflecting walls—the total energy content of the container being sufficient that the maximum entropy state involves a black hole (but not so great that the container cannot contain the hole). The phase space $\mathcal{P}$ of the contents is schematically illustrated in Fig. 9.2. The region of $\mathcal{P}$ representing configurations for which there is a black hole in the container is denoted

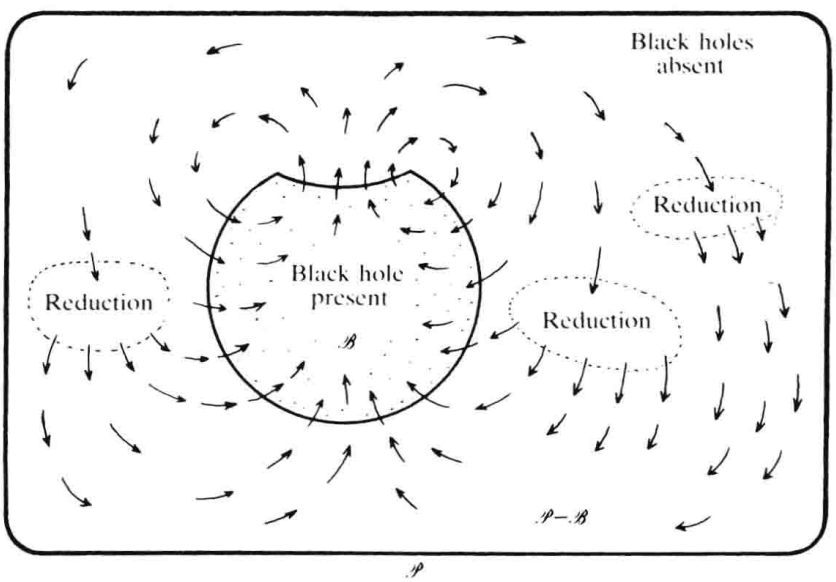


Fig. 9.2. Phase space $\mathcal{P}$ of a large isolated container with perfectly reflecting walls.

by $\mathscr{B}$. White holes (the time-reverses of black holes) are excluded since their singularities violate the Weyl curvature hypothesis. This time-asymmetric constraint on what can happen inside the container leads to some interesting consequences. However, I should make clear that not everyone would be in agreement about this picture. In particular, Hawking (1976) takes the view that black holes and white holes are to be regarded as physically *identical,* and so he would reject the hypothesis. In my own opinion this identification carries with it some severe difficulties, mainly because the classical space–time geometries of black and white holes are not at all the same, especially when one considers their modes of formation and disappearance; cf. Penrose (1979, 1982*a*). Hawking considers that this problem can be resolved if one takes the view that the geometries cannot be adequately considered as classical and must be taken to be subject to some decidedly large quantum uncertainties.)

I am being deliberately vague as to whether the phase space $\mathscr{P}$ depicted in Fig. 9.2 is to be regarded as a classical or quantum phase space (i.e. as a symplectic manifold or a Hilbert space). It is hard to see how the black hole geometries can (as of now) be treated other than classically; yet one is also concerned with the Hawking process, which is quantum mechanical, as providing the only way that black holes can disappear once they are formed. I prefer to think of the phase space as being, in reality, some sort of combination of a quantum and a classical one. But for the purposes of the present discussion I shall just remain somewhat vague about the matter. We note, however, that since the container has finite volume and a bounded energy (mass) content, $\mathscr{P}$ must, if classical, be a *compact* symplectic manifold and, if quantum mechanical, a *finite-dimensional* Hilbert space.

We recall that whereas black holes can form by means of classical collapse, they can disappear only quantum mechanically, the mass being reduced through the process of Hawking evaporation. Since according to the view to which I adhere, black holes and white holes are physically distinct objects, the classical collapse modes are not time-reverses of anything which can be achieved by means of Hawking evaporation. The implication seems to be that there are more ways for a black hole to form than there are for it to disappear. In terms of Fig. 9.2, this means that there are more routes (in terms of phase space volume or quantum mechanical state count) entering the region $\mathscr{B}$ than there are leaving it. As far as the region $\mathscr{B}$ itself is concerned the violation of Liouville's theorem or of quantum mechanical unitarity that this implies is not paradoxical because a space–time singularity is present for the states within $\mathscr{B}$. Information can be destroyed at this singularity in violation of the normal assumptions underlying Liouville's theorem and unitarity, and this accounts for the loss of phase-space volume or of quantum states

within $\mathscr{B}$. This situation occurs because the singularities are future-singularities (i.e. information absorbing singularities), all past-singularities (i.e. information creating ones) that otherwise might occur within the container having been excluded by the hypothesis on p. 138.

However, with regard to the remaining region $\mathscr{P} - \mathscr{B}$ there is a definite problem: some effective violation of Liouville's theorem or unitarity seems to be implied, yet no singularities occur for that region. More routes *leave* the region $\mathscr{P} - \mathscr{B}$ (which has finite total volume) than enter it. Somehow, phase space volume is *created* within $\mathscr{P} - \mathscr{B}$, yet the region is concerned with situations of 'ordinary' physics, where no exotic phenomena associated with the presence of black holes or space–time singularities occur. My contention is that this effective violation of Liouville's theorem or unitarity is an objective *state vector reduction*. This is the other side of the coin from our hypothesis and seems to be needed for a consistent picture. My suggestion has been that the origin of the Second Law is to be regarded as a (quantum) gravitational phenomenon (i.e. determined from space–time singularity structure) and now we appear to need the compensating idea that state vector reduction process is also a (quantum) gravitational phenomenon. For a good many years, Károlyházi has argued on quite different grounds that an objective physical reduction process should occur by virtue of quantum gravitational considerations (see Károlyházy 1966, 1974, and Chapter 8; also Komar 1969). Here we appear to have some significant independent support of this idea.

Indeed, without some spontaneous reduction taking place within the container it is difficult to make any sense of the geometry within it whatsoever. Note that, according to my picture, this reduction must (at least sometimes) be a time-asymmetric process. As a single quantum state evolves via this reduction procedure, it becomes a collection of possible alternative states, only one of which is realized in the actual world (though it could be a different one on the next 'Poincaré cycle'). We get an effective spreading of phase space volume only because if we reverse the procedure fewer alternatives are available. To understand what is involved here, imagine a simple quantum experiment. A photon is aimed at a half-silvered mirror and photocells detect whether it has been reflected or transmitted (Fig. 9.3). One state goes in; then after reduction, one of two alternative states then takes place. The initial state vector is referred to a new basis and the two constituents of this initial state with respect to this new basis are both allowed. But reverse the process in time and we find that only one of the two resulting states is reasonable and the other is 'absurd'—the reasonable one corresponding to the photon emitted from the laboratory source and the 'absurd' one, to that where the photon is ejected from the laboratory wall. It is 'absurd'

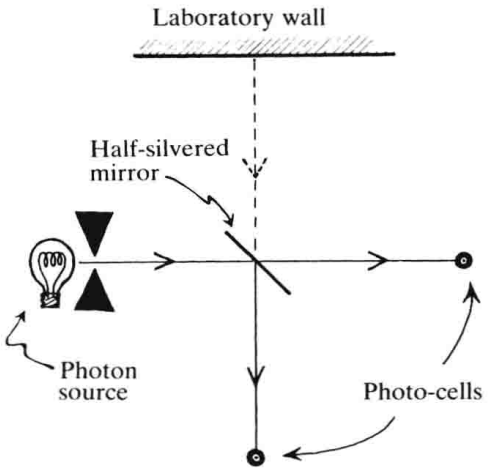


Fig. 9.3. State vector reduction in a simple case: is it time-reversible?

only for thermodynamic reasons: in ordinary time, the ejection process violates the normal progression of increasing entropy. From my own point of view, this 'absurdity' would (almost always) entail that if we evolve it backwards in time to the beginning of the universe, we get a contradiction with our hypothesis, since that is the origin of the Second Law. The reduction process and the Second Law would seem to have to be intimately connected with one another, which is a strong reason for believing that gravity must have something to do with reduction, since (on almost any point of view) it has a special role to play in connection with the Second Law.

9.3. Gravitational entropy and reduction

I shall argue that a key ingredient in the process of spontaneous state vector reduction is the concept of *gravitational entropy*. The idea is that reduction will occur when the decrease in entropy involved in the reduction effect is at least compensated by a corresponding increase in the gravitational entropy. Imagine the wave function of some particle spreading out through space, say from the decay of a radioactive atom. The geometry this produces will not have any significant Weyl curvature. But if the wave function becomes coupled to some larger system so that the different particle positions produce sufficiently clumped configurations of mass, then each individual clumped configuration may have sufficient Weyl curvature to represent a significant rise in the gravitational entropy. Under such circumstances, I would contend, reduction to one of these individual clumped geometric configurations could take place.

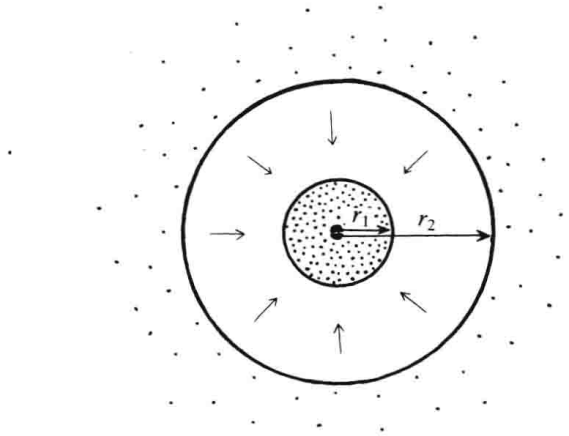


Fig. 9.4. Formation of a droplet—for the rough estimation of gravitational entropy increase.

Of course, such a suggestion requires a change in the standard rules of ordinary linear quantum evolution. But the 'rules' are themselves somewhat unclear in the context of curved space–time theory. One requires some analogue of the space- and time-translations to provide the 'natural' co-ordinate variables in terms of which quantization is to be carried out. In a general curved space–time, when such symmetries are absent, there are essential ambiguities inherent in the quantization procedure (Komar 1971). If one envisages differing space–time geometries being linearly superposed, then one must imagine that the differing 'quantization procedures' that are being adopted, for the various space–times must also be in some sense superposed. It is, to say the least, somewhat obscure what the 'standard rules' of quantum theory actually *are* when space–times of differing intrinsic geometry are superposed. So 'a change in the standard rules' for such circumstances is not such a shocking innovation as it might otherwise appear to be!

I envisage that nature feels uncomfortable about 'linearly superposing' space–time geometries which differ significantly from one another and, instead, prefers to settle for (i.e. to 'reduce' to) essentially just one of the geometries involved. We may think of this reduction as perhaps taking place via some nonlinear instability. But I would not expect that the geometries are ever quite precisely defined as classical structures. Each classical geometry would itself be some sort of average of contributing quantum states. Then, roughly speaking, the logarithm of the number of quantum states involved would provide a measure of the *gravitational entropy* that is to be associated with the classical geometry. The idea is that 'reduction' can take place when the apparent lowering of entropy

that such a process leads to is actually more than compensated by a rise in this gravitational entropy.

Though I have no clear-cut mathematical expression for this gravitational entropy for a general classical geometry, some guides to the rough form of such an expression can be suggested. In the first place, we demand agreement with the Bekenstein–Hawking formula (9.3) for a stationary black hole. But in the general case, we imagine that the entropy measure could be 'time-dependent', i.e. it would depend upon the choice of some space-like hypersurface† $\mathscr{H}$ within the space–time $\mathscr{M}$. I envisage (perhaps) some integral expression involving the Weyl curvature of $\mathscr{M}$ at $\mathscr{H}$, where the 'density' to be integrated over $\mathscr{H}$ has the rough form: integral of Weyl curvature times second integral of Weyl curvature (i.e. something at the level of 'Christoffel symbols times metric'), but where something of the nature of a positive/negative frequency split may first have to be performed. The motivation for this is that such an expression exists in the source-free weak-field (linearized) limit of the gravitational field (cf. Fierz 1940) where one is here using an estimate of 'graviton number' in a classical field as a rough guide to its entropy. The aforementioned expression provides the expectation value of the number operator.

I wish to attempt a crude estimate of this gravitational entropy in a situation of relevance to the detection of a particle in a cloud chamber or bubble chamber. When a droplet condenses in a cloud chamber, associated with its more clumped state will be a space–time geometry with a slightly higher gravitational entropy than that of the uniform distribution of uncondensed gas. It is only the gravitational contribution to the entropy that concerns us here, and one might, at first, anticipate that this contribution would be insignificantly tiny. But for an idealized spherically symmetrical uniform droplet forming (see Fig. 9.4) in which a region of radius r_2 condenses to a smaller one of radius r_1, we can imagine that a non-zero Weyl curvature occurs only between the radii r_1 and r_2, the (linearized) Schwarzschild solution holding there (assuming, for crude calculational purposes only, that a vacuum is left between r_1 and r_2). The entropy 'density' to be integrated thus occurs between the radii and has

† The view has been expressed by Hawking and his collaborators that only when there is an absolute event horizon should there be a non-zero gravitational entropy. However, because of this dependence upon data at a (space-like) hypersurface it is not a clear-cut matter whether such a horizon exists or not. The black hole metrics are not, at $\mathscr{H}$, clearly distinguished from those that are not black holes. (If an apparent horizon is present on $\mathscr{H}$ then, with the weak energy condition holding, a horizon *will* exist. But this is not the correct *criterion* for a horizon. In any case, the 'turning on' of an entropy only when some inequality is satisfied seems to me to be quite inappropriate.)

the rough form of a constant times

$$\frac{m}{r^2}\times\frac{m}{r}, \tag{9.6}$$

where m is the mass contained within the (smaller) radius, so we anticipate a total gravitational entropy of something like

$$(\text{numerical constant})\times\left(\frac{kG}{\hbar c}\right)m^2\log\left(\frac{r_2}{r_1}\right). \tag{9.7}$$

This expression diverges for $r_1\to 0$, but we expect a cut-off at the Schwarzschild radius $2mGc^{-2}$. The formula (9.7) is presumably only approximate and should go over to the Bekenstein–Hawking value (9.3) at roughly this cut-off radius.

Despite these various uncertainties and the obvious crudeness of eqn (9.7), we can use eqn (9.7) to give us a very rough estimate of the gravitational entropy in a cloud chamber droplet. Similarly we can use it for a bubble in a bubble chamber, where in this case we regard it as resulting from a condensation of negative mass over and above the positive-mass background. Because m occurs squared in eqn (9.7), the entropy contribution is still positive. (The formation of a bubble is still a *positive* clumping.) We find that if we need, say, an entropy of $100k$ to reduce the wave function of a particle which expands spherically symmetrically from a radioactive source in the laboratory to one localized in a track in the cloud chamber or bubble chamber, then either one rather large droplet or bubble is needed, of perhaps a millimetre in diameter, or, more appropriately, a string of somewhat smaller ones. This all seems rather on the large side, and one might have anticipated that a physically real state vector reduction ought to have already taken place before this. But the results are *not absurdly* unreasonable—as they could very easily have turned out to be! Had the logarithmic expression (9.7) been, instead, something like the $Gm^2\,(r_1^{-1}-r_2^{-1})$ that one obtains for the gravitational *self-energy* contribution, then there would have been no chance whatever of being able to produce a significant value on the scale of a laboratory experiment.

I regard these results as providing some general encouragement for the point of view that I am promoting. But clearly we need a very considerable improvement in the gravitational entropy expression. The 'correct' expression (assuming that there is one) must clearly be of a non-local character. This is already true of gravitational *energy*, but it does not prevent the possibility of a precise formula at the quasi-local level, (cf. Penrose 1982*b*; Penrose and Rindler 1985). For entropy the non-locality may be one stage more 'nebulous' since there is one further degree of integration involved in the weak-field limit expression. The

hope is that this non-locality can be matched with the kind of non-locality that quantum phenomena so frequently exhibit (cf. Chapter 4). This is yet a matter of speculation, but it opens up a number of intriguing possibilities.

Acknowledgements

I am grateful to many friends and colleagues for enlightening discussions, most particularly, to I. Bialynicki-Birula, P. Pearle, I. C. Percival, and D. Deutsch.

Postscript

In subsequent discussions, R. L. Wald has more-or-less persuaded me that my tentative expression (9.7) should not be interpreted as an entropy in the ordinary sense. I think that it is more correctly just a measure of the increase in *graviton number* and that the whole discussion should perhaps be rephrased in such terms. Recently, A. Ashtekar has informed me of a calculation he has performed, indicating that the numerical constant in expression (9.7) should be something like

$$768\pi^3 \doteqdot 2.4 \times 10^4.$$

This is gratifyingly large: it means that the droplet or bubble size, for reduction to take place, would be considerably smaller than I had previously estimated.

I am now suggesting that reduction takes place whenever *one* new ('longitudinal') graviton is produced.

References

Everett, H. (1957). *Rev. Mod. Phys.* **29,** 454.
Fierz, M. (1940). *Helv. Phys. Acta* **13,** 45.
Hawking, S. W. (1976). *Phys. Rev.* D**13,** 191; **14,** 2460.
Károlyházy, F. (1966). *Nuovo Cimento* A**42,** 390.
——, F. (1974). Gravitation and quantum mechanics of macroscopic bodies. *Magyar Fizikai Polyoirat* **12,** 24 (Thesis, in Hungarian).
Komar, A. B. (1969). *Internatl. J. Theor. Phys.* **2,** 257.
—— (1971). Semantic foundation of the quantization program. In *Studies in the foundations, methodology and philosophy of science,* Vol. 4, *Problems in the foundations of physics* (ed. M. Bunge). Springer, Berlin.
Penrose, R. (1978). Singularities of space–time. In *Theoretical principles and astrophysics and relativity* (ed. M. R. Lebovitz, W. H. Reid, and P. O. Vandervoot). Chicago University Press, Chicago.
—— (1979). Singularities and time-asymmetry. In *General relativity—An Einstein centenary survey.* (ed. S. W. Hawking and W. Israel). Cambridge University Press, Cambridge.
—— (1981). Time asymmetry and quantum gravity. In *Quantum gravity* 2. Oxford University Press, Oxford.

—— (1982*a*). Some remarks on gravity and quantum mechanics. In *Quantum structure of space and time* (ed. M. J. Duff and C. J. Isham). Cambridge University Press, Cambridge.
—— (1982*b*). *Proc. R. Soc. Lond.* **A381,** 53.
—— (1986). On the origins of twistor theory. In *Gravitation and geometry,* I. Robinson 60th Birthday volume, (ed. W. Rindler and A. Trautman) Bibliopolis, Naples.
—— and Rindler, W. (1985). *Spinors and space–time, Vol. 2; Spinor and twistor methods in space–time geometry.* Cambridge University Press, Cambridge.
Von Neumann, J. (1932). *Mathematical foundations of quantum mechanics* (Princeton University Press, 1955) Ch. VI.

167
BIG BANGS, BLACK HOLES AND 'TIME'S ARROW'

The nature of time, 36–62. Blackwell, Oxford, 1988

Big Bangs, Black Holes and 'Time's Arrow'

ROGER PENROSE

What I have to say relates to the beginning and end of the universe; but also to perfectly commonplace events, such as what happens when a cup of coffee is stirred. With regard to the beginning of the universe, some of my remarks may be a little unfashionable. That is to say, I shall not talk about the current ideas of 'inflationary universes' (with which I am somewhat out of sympathy). What I shall be concerned with is the seeming paradox that the laws of physics, or such of them as we understand, appear to be reversible in time, whereas we are constantly confronted with large-scale phenomena that are manifestly not reversible. Think, for example, of atoms colliding with each other. The laws which govern the behaviour of such small particles are completely reversible in time, yet the macroscopic objects which are composed of vast numbers of these tiny objects behave in ways which are patently irreversible. And that applies to such things as people and so on. Take a film of almost any human activity, run it backwards, and what you see corresponds to nothing that one would, or normally could, encounter in real life.

There is, of course, a kind of pat answer one can make to this, and that is to appeal to the second law of thermodynamics. This law says that there is a thing called *entropy* and that this entropy goes up as time goes on. We postulate that. Now that *is* an irreversible law. But it's a law which applies to large systems of particles, not to individual particles. It says, roughly speaking, that if things are peculiar early

on, in certain manifest ways, then they will get less and less peculiar as time passes. A typical example is a box containing gas, which starts off with all the gas up in one corner. If you leave it, the gas will spread itself uniformly all over the box. When the gas is all in one corner, the system is in a low entropy state; and when, by contrast, the gas is uniformly distributed with a certain very specific distribution of velocities and so on, it is in a high entropy state. Thus it evolves from the low entropy to the high entropy state.

Actually there is nothing very remarkable about the fact that the entropy of a system goes up, because all that really means is that if you start a system off in a very special state and you let it go, it is likely to end up in a less special state. It may be useful here to introduce the notion of what is called *phase space*. Phase space is an imaginary space with a large number of dimensions, which is used to describe a physical system. The state of the system, at any given time, will correspond to a single point in that space. (In classical mechanics, a system of N point particles can be represented by a space of $6N$ dimensions, six for each particle, corresponding to its position and its momentum with respect to the x, y and z axes, respectively.) If we consider a phase space for our box containing gas, then there will be certain regions of this space which are manifestly peculiar — the gas being all in one corner, for instance. Such regions of phase space will be very small. But there are other states which look very much the same as each other, such as the states in which the gas is pretty uniformly spread out all over the interior of the box. These states will jointly occupy a much larger region of the phase space. Clearly, if you start the state off in a manifestly peculiar region of phase space, and you just let the state evolve, so that the phase-space point wanders around somehow, the chances are that it's going to get into a large region.

That is more or less what the second law of thermodynamics says: namely, that if you start in some tiny region of phase space that has been specially singled out, or something like that, then you will soon find yourself in a much larger region. This thing called entropy, which is usually denoted by the letter S, is, roughly speaking, the logarithm of the size of the region of phase space in question — a region singled out in terms of its manifest macroscopic properties. One thing that should be made clear is that there is an absolutely ridiculous difference in the sizes of these different phase-space volumes. (This is one reason why one takes the logarithm. It's not the main reason, but one advantage of taking logarithms is that the

differences then don't look so extraordinarily big.) If, on the other hand, you are just told the entropy values corresponding to two regions of phase space, then you have to take the opposite of logarithms, namely exponentials (that is, raise to the power of those numbers) to find out the relative sizes of the phase-space regions.

There's nothing at all remarkable about the fact that the entropy goes up in the future. Nevertheless, there *is* a very peculiar thing about the second law of thermodynamics. It actually is a most extraordinary physical fact, and what's extraordinary about it is that if you look into the *past* instead, you find that things don't behave in this apparently 'probable' way. Let's take Newtonian physics; forget about quantum mechanics for the moment. If you take standard Newtonian physics, the laws are time-reversible, and you can just as well evolve into the past as you can into the future. Both are equally allowable. So let's go into the past. If we evolve the state into the past, we find that, because the second law says the entropy's going up in the normal direction, the entropy is going down as we go into the past. You start with a point in some region in phase space, and you find that as you evolve into the past, the point goes into more and more peculiar regions as time proceeds (backwards).

This may not seem particularly remarkable, but only because we are used to thinking of things evolving forwards in time. Didn't I say that the second law of thermodynamics is a time-irreversible law? Well yes. It is, if it's stated in the simple form: entropy always increases in the direction earlier to later. But when this is explained along the lines sketched above, as simply a statistical consequence of the ordinary reversible laws of physics, as applied to very large aggregates of particles, the second law seems to be being deduced from something time-reversible! Suppose you take a closed system (free, that is, from external interference), examine it at time T, and find that its entropy is less than the maximum possible for the system. Then statistical mechanics tells you that, if you examine it at any time later than T, the chances are that it will be in a state of higher entropy. But, in the absence of other information, save what the state was like at time T, the very same argument says that if you examine it at any earlier time, it is likewise probable that it will be found to be in a state of higher entropy! Why, then, don't we actually find (or expect to find) such behaviour in the majority of cases? The answer is that we interpret the state of relatively low entropy as being a consequence of it having been started up in a state of even lower entropy, not of a random fluctuation away from a state of high

entropy. As the jargon has it, the temporal asymmetry derives not from the evolution laws so much as from the assumed *boundary conditions* in the past: the assumptions we make about how things started off.

So the puzzle is: why was the entropy so low in the old days? Well, when were the old days? Clearly, entropy is increasing now, and has been for a very long time. The question, then, is how far back one is entitled to extrapolate that trend. Well, physicists usually just assume the second law. They don't think that at a certain time the rules changed, and before that entropy was going the other way. So we must take it that it just keeps on going down as we go further and further back in time. It looks, therefore, as though we might as well go right on back until we get to the beginning: time zero. Well, there is supposed to *be* a beginning on the modern view of cosmology — that is, the big bang. And in fact there is quite a lot of evidence from various quarters that there was a highly condensed state in the early stages of the universe. There is evidence both from theory and from observation that there was some such thing as the big bang, that the universe was originally concentrated in some very tiny region. The entropy must presumably have been at its lowest value at the big bang.

That's fine. But, on the other hand, there is a sort of paradox involved here. I said there was quite a lot of evidence in favour of the big bang. Some of that evidence comes from detailed calculations that astrophysicists carry out, when they think about the early stages of the universe. And one of the prime ingredients which goes into these detailed calculations is that the universe, in its very beginnings, was in a state of *thermal equilibrium*. Now, thermal equilibrium — what does that mean? That means that you're in a state of maximum entropy. Well you don't have to be a mathematician or physicist to realize that there's something a little bit puzzling here. The entropy has gone down and down and down in the past, until it reached its maximum value!

I've been puzzled by this for quite a long time — very many years. I've often mentioned it to my astrophysicist friends and I've got this kind of strange look usually. I felt that I must have said something stupid. People obviously thought that I'd somehow missed something — some essential ingredient in this whole argument. Well I suppose the essential ingredient that I'm supposed to have missed here was the fact that the universe is expanding. The idea seemed to be that one starts from this very small state, the universe gets bigger and for

some reason the entropy that's allowed in the early stages is very, very tiny. The universe was using up all the available randomness, if you like, and being as random as it could, but since the universe was so small, the possible maximum entropy was much lower than it is in the universe today. Well, that perhaps kept me quiet for a little bit; I'm not sure. But, after thinking about that for a while, I realized that it doesn't make any sense. It doesn't make any sense for a perfectly obvious reason. And one thing that worries me about this whole discussion, actually, is the fact that it only involves ingredients which are perfectly obvious. How is it that one can come to a conclusion with which everybody seems to disagree, when you're only putting in things which are perfectly obvious?

So why doesn't this make sense? Well, it might be the case — people don't know for sure — that the universe is closed, finite, a kind of three-dimensional analogue of the surface of a sphere. That's one possibility. I should say that the evidence, for what it's worth, is probably slightly in the other direction. But that's not the point, because clearly it *might* be the case that the universe is closed. So let's suppose that the universe *is* closed. Then what will happen? It will reach its maximum expansion, turn around, come back again and will then finally end up in what's sometimes referred to as the *big crunch*.

Now the suggestion we're currently exploring — the standard response to my puzzle — is that the universe indeed started in a state of thermal equilibrium; but that that still corresponded to a very low total entropy, because the small size of the universe placed a very low ceiling on the maximum entropy it could have. For a system to be in a state of thermal equilibrium just means its having the highest entropy possible, given its size. So what happens, on this view, as the universe expands? Well very early on — so the story goes — the universe is expanding very rapidly, so rapidly that the expansion rate exceeds the rate at which the system is able to evolve back to a state of equilibrium, in obedience to the second law. So at first the universe is getting further and further away from thermal equilibrium. As the rate of expansion slows, however, the rate at which the universe approaches thermal equilibrium once more exceeds the expansion rate. Well, that may initially strike one as quite reasonable. But if this general line of thought is correct, what is supposed to happen to the entropy when, as we're now assuming, the universe begins to contract? On the one hand, the second law of thermodynamics says the entropy must keep right on going up. But,

on the other hand, the view I've just sketched clearly entails that it *can't* go on doing that indefinitely. For as the universe continues to contract, the maximum possible entropy will get progressively smaller: and you can't get a quart of entropy into a pint pot, so to speak. We appear to have a contradiction.

Well, there is a view on this subject which I should perhaps mention. But I don't want to spend too much time on it, because I don't believe it myself and I think one can present a strong argument against it. The suggestion is that perhaps somewhere in the middle of its history, the universe actually reaches its maximum entropy, which then starts going down again, as the universe contracts. This may sound plausible. But you have to worry about what it actually entails. Current thinking has it that it is entropy that determines the arrow of time. It is highly plausible that the macroscopic asymmetry of time is grounded in the second law of thermodynamics; that it is that which ultimately explains such things as why we can remember the past but not the future. If that's correct, then what makes us think of one direction in time as 'earlier' and the other as 'later' is that the one direction corresponds to decreasing and the other to increasing entropy. So if entropy were to stop increasing and start decreasing again, it would seem to follow that time itself would come to have a different directional sense. Things after the turnaround would behave as though the future lay in the direction we think of as the past. And that gives rise to all sorts of problems. One problem is that the universe starts to recollapse, for instance, before you see it recollapse and the stars are all coming towards you. So what does the light do when it starts going the wrong way? What exactly is supposed to happen to light rays that span the transition between the period at which entropy is increasing and that at which it is decreasing? Will light rays diverge from a source prior to the turnaround, and then miraculously reconverge on some point after the turnaround — a point which, from the post-turnaround perspective, constitutes their source?

We also have trouble with various paradoxes. Suppose, for example, you have people living after the turnaround, in the contraction era. We can call them the *contraction people* (though they, of course, will think of us as the contraction people, and themselves as expansion people). They, presumably, will remember things that happened in what, from our perspective, is their future. If they were to be told things that had happened in what we would regard as their past, this would create a paradox. Because what

would then prevent them from choosing to do something different? But now *we* are in what they regard as their future. So why couldn't we send the contraction people a message, telling them about something they would regard as in the future: something, moreover, which it would seem to be in their power to do something about? What we could do, in principle, is put a lot of information in a sealed container and simply let it loose in the universe. If current estimates about the size of the universe are correct, there wouldn't be much difficulty, in principle, in building some box that would last through to the contraction phase, on the assumption that the universe will eventually start to contract. There is then no obvious reason why the contraction people shouldn't come across this container, find out what was going to happen in their future and do something to prevent it happening. Suppose, for example, they were to send out a space probe of their own on such a course that it would rendezvous with the container our side of the turnaround and destroy it, thereby preventing them from ever receiving the container in the first place? Actually, things are not quite as simple as I've made them sound. But there are, in principle, very severe difficulties with this kind of picture.

But you might say to me: 'All right, then, perhaps the universe isn't closed.' And, after all, as I said, the evidence, for what it's worth, is perhaps somewhat on the other side, suggesting that the universe may well continue to expand indefinitely — in which case these problems won't arise. Well, that doesn't really get you out of the difficulty. And this is where the black hole part of the title comes in.

At this point it will be useful to draw a diagram. It's quite helpful, sometimes, to draw pictures of the whole universe on a blackboard or on the back of an envelope or something like that. Here we can make use of the sort of devices employed in certain map projections, where we don't mind distorting distances, but don't wish to distort angles. What I am talking about here are *space—time* diagrams; and the idea is to avoid distorting the *light cones* too much. A light cone is an imaginary surface, associated with a point in space—time, comprising the paths of all possible light rays that pass through that point. Time progresses from the lower to the upper part of the diagram, and a particle, or any other object, is represented by a curve referred to as its *world-line*: this gives the history of that particle as time progresses. The light cone is the important ingredient here, since it shows you how signals can propagate. Particles necessarily have

their paths constrained to lie within the light cones: if you take any point in space—time, and take the light cone associated with it, then any world-line that passes through that point must lie within the corresponding light cone. This is just another way of saying that no particle can travel faster than the speed of light, or more precisely, that no particle can exceed the *local* light velocity. Now in general relativity, you may have light cones drawn all higgledy-piggledy and so on (provided there is a smooth transition from the light cone you get at any one point to that which you get at any other). But the essential thing is that if you consider the local observers travelling around on their world-lines (if that's the right way of talking — strictly speaking they *are* their world-lines), then these world-lines have to be constrained, wherever they are, to be within the light cones. So if you think of it locally, time will be going in the direction from the backward to the forward part of the light cone. No-one is allowed to cross the boundary of his own light cone. So that's the sort of picture one has. And if you imagine these cones being drawn on a rubber sheet, then general relativity allows you to stretch and bend them around locally, if you like. What you get will still obey the same fundamental laws.

What you can do, though, is draw pictures where these light cones, rather than looking distorted or leaning over at crazy angles and so on, all sit up in a fairly orderly fashion. Let's start with the more standard type of space—time diagram. We can, for example, represent the standard cosmological models in quite a useful way. The closed universe that I was talking about just now starts off with a big bang — at least, that's the way people usually think of it. The big bang seems to be a singularity. That's where all the curvatures go to infinity and so on. So it can be represented as a point at the bottom, while the big crunch is represented as a corresponding point at the top. It's useful, here, to throw away two dimensions and think of the universe as having just one space and one time dimension for the purposes of visualization. So we can think of it as the surface of a sausage of some sort, with space going around the circumference and time running lengthwise from the bottom to the top (see fig. 4.1).

That's one way of representing the history of a closed universe. Here the light cones are all arranged lengthwise along the surface. But then they have a habit of getting rather squashed when you approach either end. And to see what's happening in detail in the vicinity of the big bang and the big crunch is consequently rather difficult with this picture. What's useful, therefore, is to stretch this

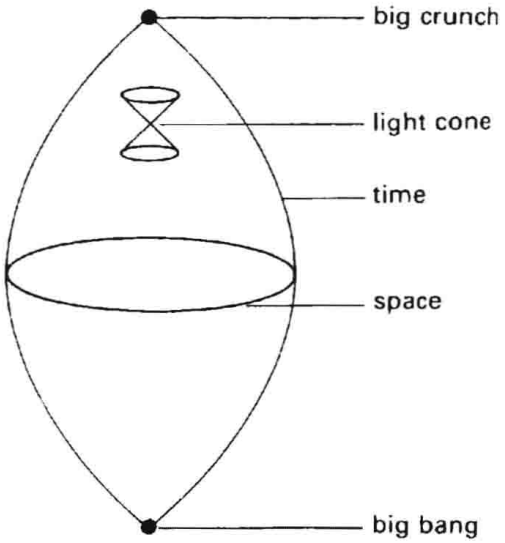


Fig. 4.1

model at the bottom and the top, and open it out like a map. When it's unfolded it looks something like fig. 4.2. This represents the whole universe. The lower and upper edges correspond to the big bang and the big crunch, respectively. And the light cones are now drawn all the time at 45°, no longer becoming distorted as we approach the two singularities. That's the advantage of this form of diagram; and it makes for much clearer representation of the important features. All the standard closed universes postulated by cosmologists may be depicted in this kind of way.

Imagine now a particle emerging from the big bang, and somebody sitting on it looking out at the universe. One thing you will find is that, as time goes on, more and more of the universe is encompassed in his field of vision. At any one time, only a certain set of particles will be visible to the observer. The boundary of this set is called his *particle horizon*. At any given time he can, in principle, only see a certain region of the universe. A particle that lies on the far side of his particle horizon is one that has not yet come into view. But if he waits a bit longer (always assuming he lives long enough) he will eventually see it. This is the feature of a particle horizon. Now you'll notice that the big bang is, in fig. 4.2, stretched out into a whole space-like region. This brings me to something that often worries people: something which starts off at point A, say, can have, initially, no influence on something that starts off at point B, because A lies

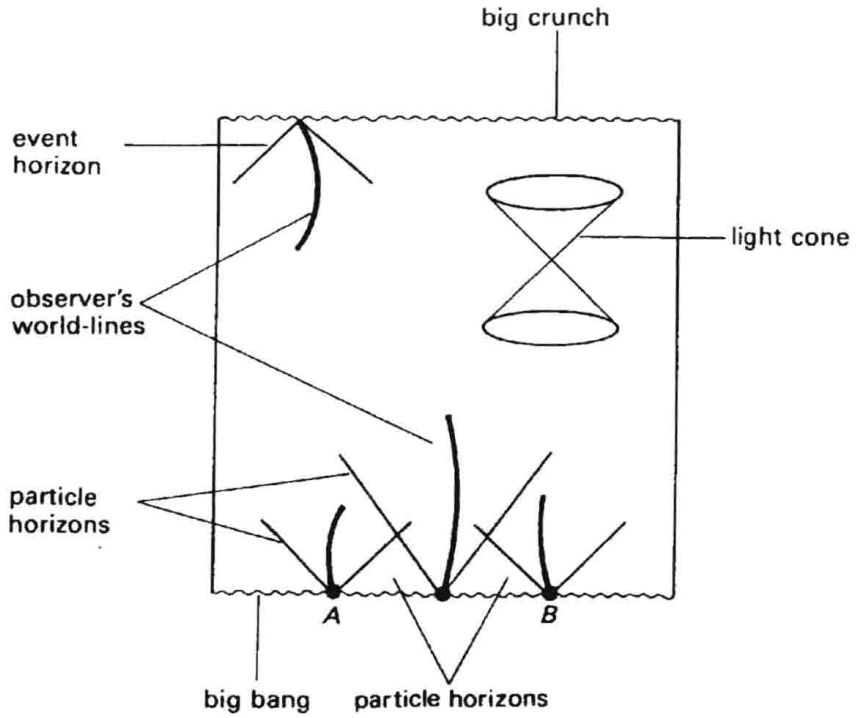


Fig. 4.2

outside the light cone, and hence outside the particle horizon, of B. A lot of difficulties that people have with the standard cosmological pictures result from these so-called *horizon problems* in cosmology. There is, for example, a lot of activity these days devoted to what are called *inflationary models* of the universe. And one of the main aims of these models is to try to remove this seeming difficulty that one has in thinking about the early stages of the universe, and to try to overcome these horizon problems. But I don't want to say anything more about these inflationary models here, other than pointing out that, even though these ideas of mine were developed before inflationary models became fashionable or even known about, the inflationary models don't in any significant way affect the present argument.

There is also another kind of horizon that I should mention (also shown in fig. 4.2). If you have an observer approaching the big crunch, there are some events which are forever invisible to him: these are said to lie beyond his *event horizon*. By the time this observer is finally crushed out of existence, these events will still not have been observed. Had the observer chosen to do something else at

a sufficiently early stage, so that his world-line went in a different direction, then he would have been able to observe some of these events. But then there would have been other events which he would have been completely unable to observe: events which, as things are, he can observe.

Event horizons thus stand to the big crunch as particle horizons stand to the big bang. The possession of such horizons is a standard feature of these models. I mention them only because they keep coming up and because I wanted to draw you a picture of a black hole in this kind of scheme. There will originally have been a star, galaxy or something, which collapsed under its own weight, to the point where its own gravity became sufficiently intense to overcome all opposing forces. This collapsing matter is shown at the bottom of fig. 4.3. In the middle there is a somewhat unpleasant singularity. Now the point is that the light cones are, so to speak, pulled in by the gravitational field of the collapsed matter. The event horizon represents the point beyond which it is impossible for light to escape: points on the horizon thus have their light cones tangential to the horizon. Once inside the region bounded by the event horizon, an observer cannot escape, because to get back across the horizon it would be necessary for his world-line to pass outside his own light cone: for it to go in what is called a *space-like* direction. And this is not allowed. So all he can do is go from the outside to the inside. He cannot turn round and go back again without exceeding the local light velocity.

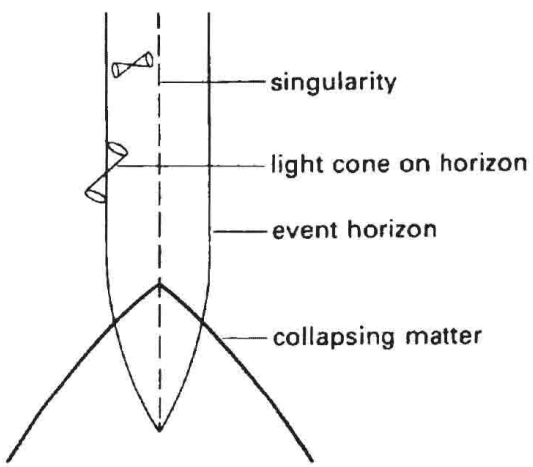


Fig. 4.3

So that's what a black hole looks like, drawn in the standard way. But that is not always the most useful way to draw it. In fig. 4.3 the singularity is shown as a vertical line, and the light cones are all falling in towards it. If, in accordance with our new type of diagram, I turn the light cones round so they look upright, then you find that the singularity becomes horizontal (see fig. 4.4). The event horizon is now shown as a line at 45° to the vertical. Then out to the right-hand side you've got infinity, which can be included if you like. I shouldn't worry too much about that. And on the left we have the axis of symmetry. You should think of the whole thing rotated round that axis. In reality, it's centred on the dotted line.

Now suppose we have a brave astronaut who decides to fly his spacecraft into a black hole in order to carry out various experiments in connection with the second law of thermodynamics. He keeps taking glasses of water and smashes them on the floor. Several rather crude experiments, I'm afraid. But then he happens to be a rather crude astronaut!

I'm still addressing this original worry of mine. Suppose it were true that the reason that the entropy is going up is merely that the universe is expanding. Well here in the black hole, you have locally a universe which is collapsing about you. The future of our exceedingly brave astronaut is dominated by the singularity. Now suppose you're taking the view that whatever happens to the universe is governed

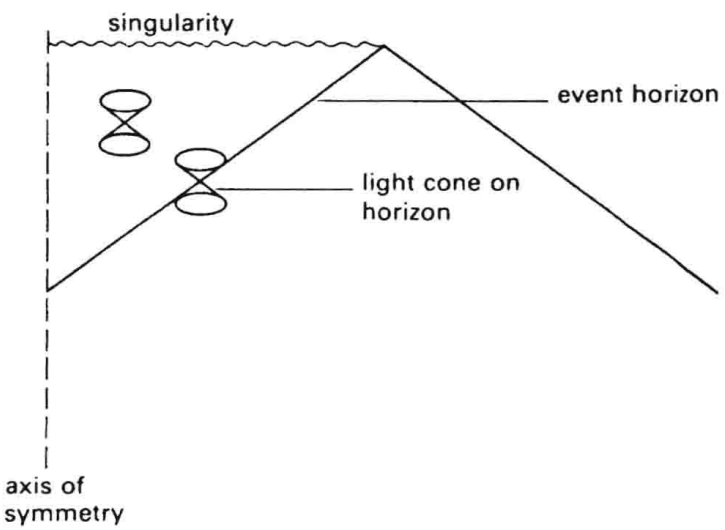


Fig. 4.4

by its boundary conditions. Suppose, further, that you're not going to be biased as regards the directions forward into the future or back into the past; and seeing that the laws of physics locally are time-symmetric, there's no reason to be biased that way. That means, then, that whatever happens to that astronaut's entropy should be governed by the singularity. His universe is collapsing completely. According to the hypothetical view we're now considering, entropy would go down as the singularity is approached. We should therefore look at things the other way around, for the purposes of applying the second law. Things inside the horizon would behave as though, from their own point of view, this region of the universe was expanding. Regarded from this time-reversed perspective, there is no information coming from anywhere but the direction of the singularity: what we think of as an event horizon acts as a particle horizon. Somehow, according to this view, the astronaut seems to find entropy going into reverse as he crosses into the interior of the black hole.

Well, I don't believe it. And I don't think many people would. In principle this is an experiment that could actually be carried out, because there is very possibly a large black hole in the centre of our galaxy. Of course it's not something one would do lightly: it would cost a lot of money for one thing. But in principle, there is no reason why one shouldn't travel to the galactic centre and enter this black hole. It would leave a certain amount of time, perhaps a few hours. At any rate, for *certain* galaxies, it would be a number of days. So you would then certainly have sufficient time to conduct some experiments. But I don't think many people really believe that the direction of the increase of entropy would suddenly turn around.

In fact, even when you're just outside the event horizon, the future is still dominated by the singularity. When the astronaut is only just outside, then from the time-reversed perspective, only very little information is coming from outside. So you would hardly expect things *suddenly* to flip over as you crossed the horizon: you'd expect a little more warning. Odd things — such as self-assembling broken water glasses — should already have started happening before the astronaut actually reached the horizon. They should, that is, if you believe this point of view. So our astronaut ought in principle to be able to demonstrate these things and return to tell the tale. (Actually falling into a black hole is not, after all, the greatest way to die.) Perhaps he could bring films back with him to prove it (although people might be sceptical as to which way the film had been running). But, as I say, I don't really think such an entropy reversal would

take place. I think the entropy would go right on going up, and would continue to do so as the astronaut crossed the event horizon. There's no reason why it shouldn't and every reason to believe that it would.

So what's wrong with the original argument which suggests that there is a low entropy ceiling for a small universe? What's wrong, I believe, is the assumption that our state of seeming thermal equilibrium actually represents a state of maximum entropy. Not only was the universe *not*, in my view, in a state of maximum entropy at the big bang; I think it was in a state of quite absurdly low entropy relative to what it 'might' have been.

The question to consider is what actually happens to the universe, in detail, as it, or part of it, eventually stops expanding and begins to collapse. What sort of end state do you get? Well, you get something which isn't remotely like the state at the start. In fact there already exists a reasonable-looking picture of what the universe would be like in the last stages of the collapse. First the galactic centres collapse, sucking in surrounding stars and gas. And as they start coming together, these galaxies congeal with each other. Their black holes fuse to form a larger and larger congealed mass. The singularity at the end is a great mess. It's nothing like the singularity at the beginning, which seems to have been extraordinarily uniform and symmetrical.

The initial singularity is sometimes said to be a place where one hasn't the remotest idea what's going on in the physics. Well in one sense that's fair enough. But on the other hand, there's a tremendous amount of structure in the big-bang singularity. It's not at all, in its geometric form, a generic thing. By contrast, one expects the final state to be some very general type of singularity. And there's a great deal of difference. In fact, one can make some kind of estimate as to the difference of entropy owing to these congealing black holes in the end, and compare it with more everyday entropy values and with the entropy at the beginning. I'll just give you a few figures. Let's think about a closed universe. And let's suppose it to contain 10^{80} baryons: protons and neutrons that is. Why do I take that figure? Well, most people take this figure. The reason, I think, is that Eddington at one time produced a marvellous number which was supposed to be the exact number of protons in the universe. And although nobody any more believes Eddington's theory, they still use his figure for the number of baryons in the universe. That's the only reason I can think of, because there isn't really much evidence to support it. We do

know that it is at least that sort of size. But it might be much huger. If it's bigger, then things are even more extreme than I am about to suggest. So if you like, I'm taking a very conservative view.

The first line in table 4.1 gives the entropy associated with the big bang. I have put that down because, in the discussion about the early universe, there is a figure which is sometimes bandied about, with good reason, which is either 10^8 or 10^9 — something of that order. This is supposed to be the observed entropy per baryon (using units for which Boltzmann's constant is taken as unity). Now the observed value is not necessarily the actual entropy value of the big bang. It is the value that one obtains from the *background radiation*. There is this radiation out in space, which is, as it were, left over from the hot big bang. A lot of the evidence for the existence of the big bang comes from this background radiation. Now there are a great many photons for every baryon. And since the photons have lots of degrees of freedom, that gives an entropy per baryon value, which is, roughly speaking, the number of photons that there are for every baryon. (Recall that in connection with the discussion of the dimension of phase space it was noted that any particle — and this includes photons as well as baryons — has six degrees of freedom. These can be described as three for its position, at any one time, two for its direction of motion and one for its energy. For N particles there would be $6N$ degrees of freedom. For randomly moving independent particles the entropy would be essentially equal to the number of degrees of freedom. So with 10^8 times as many photons as baryons, the entropy would be dominated by the photon number, and therefore about 10^8 times the baryon number.)

Table 4.1

	S/B	S
'Observed'	10^8	10^{88}
1 $M_\odot$bh	10^{20}	10^{100}
10^6 $M_\odot$bh in 10^{11} $M_\odot$galaxies	10^{21}	10^{101}
10^{11} $M_\odot$bh	10^{31}	10^{111}
Collapsed universe	10^{43}	10^{123}

S, total entropy; S/B, entropy per baryon; $M_\odot$, one solar mass; bh, black hole(s)

People have been trying to explain why this 'observed' figure should be so huge — because it's much larger than the sorts of entropies one finds in ordinary physical processes. I've also put down the total entropy, which is quite useful. The total entropy is 10^{88}. But is that figure really so big? That's what I want to ask. Let's think a little bit about what is the right way to look at this. I talked earlier about phase space. Figure 4.5 is supposed to be a phase space for the whole universe. (I'm afraid you'll have to use your imagination here.) Every point in this picture represents a possible universe. It doesn't make much difference what stage of the universe's evolution you consider. But if you like, you could think of it as the way the universe started off. Imagine the poor Creator trying to start the universe off in a way that would result in something like what we see. Then it's a question of how carefully the Creator has to aim in order to hit a point which would give us a universe looking something like the one we actually find ourselves in. This is the question I want to address next.

Now I shall put it to you that this is actually a ridiculously small region to aim for. That's the key thing. I said earlier that my argument was going to be based mainly on things that were perfectly obvious. But in fact the thing that I want to say next is by no means

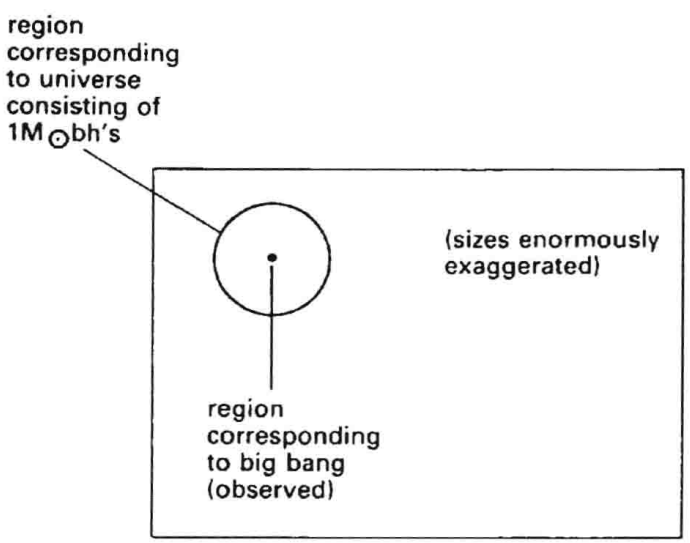


Fig. 4.5

obvious. Bekenstein and Hawking (see Further Reading) have given the following formula for the entropy S_{bh} of a black hole:

$$S_{bh} = \frac{kAc^3}{4\hbar G}$$

This is a very beautiful formula. Most of the symbols in it are actually irrelevant to what I want to say. The only things on the right that are relevant are A and perhaps 4. k is just a constant: Boltzmann's constant: c is the velocity of light, $\hbar$ is Planck's constant divided by 2π and G is Newton's gravitational constant. If you're a relativist like me, you simply put all those things equal to one. You can get away with that (though you have to observe a certain amount of caution: you can't put 4 equal to one!). But still, the essential formula is $A/4$. A is the area of the black hole, the area of the surface of the event horizon. If you know that area, this formula tells you the entropy that should be associated with the geometrical configuration of the black hole. There are various reasons for believing in it, which I don't want to go into; they derive from different points of view. Bekenstein approached the matter by way of a thought experiment, having to do with what would happen if you lowered things into a black hole. Hawking's approach was based on completely different considerations which arise when you do quantum-field theory in curved spaces. And, lo and behold, these two approaches came together. Bekenstein managed to get the A and Hawking added the 4 (Hawking's achievement was by no means the lesser one).

Anyway, it's a marvellous formula. Now let's put some figures into it. When you put the figures in you can actually work out how much entropy there ought to be in a black hole. Let's suppose, first, that every star in the universe was not a star like the Sun, but a black hole, with a mass equivalent to that of the Sun: one solar mass (symbolized as $M_\odot$). The entropy we then get per baryon for these black holes is 10^{20} — huger, much huger than the figure we started with, by an enormous amount. So, going back to our earlier question, a Creator aiming at random would have been much more likely to hit this region of phase space than the region corresponding to the actual universe (at the big bang). It's a ridiculously huger region. I have shown these as two spots on the phase-space diagram (fig. 4.5). But in fact I've had to exaggerate grossly their sizes, particularly that of the smaller one: drawn to scale it would be much smaller than microscopically small, by comparison. It's much easier to hit the

larger region (with the black holes) than the smaller one (the 'observed' one). I'll tell you by how much. If the whole universe were built out of these solar mass black holes we'd have a total entropy of 10^{100}. This means that it's $10^{10^{100}}$ times as likely that the larger region would be hit than the smaller one!

But you'll retort, of course, that the universe isn't made up of solar mass black holes. That's true. But on the other hand, there might very well be lots of 10^6 solar mass black holes in 10^{11} solar mass galaxies. These would be very large black holes, by ordinary standards. And you gain in entropy by having large black holes. Many galaxies probably do have large black holes at their centres, even though there probably isn't a very large one at the centre of our own galaxy. Having a fairly sizeable black hole in the centre more than compensates for the rest of the matter being in the form of stars. So this would correspond to a larger region still of phase space. It would have been far easier for the Creator to have hit *this* region. Well, this may still be fairly unrealistic. The actual present entropy could even be a lot bigger than this. But suppose it's smaller. I should say that that doesn't make the slightest difference. (In fact in my own calculations, when I first pursued this line of thought, I was out by a factor of nearly $10^{10^{123}}$!) Eventually, a lot of the material in this galaxy will collapse into that black hole in the centre, giving you an entropy per baryon of 10^{31}. All right. What about the very end when the whole universe has collapsed? Well, you can estimate. The black-hole calculations don't directly apply, but you can imagine that there is a larger and larger clumping of the black holes. And, for a collapsed universe, you're going to get an entropy per baryon of something like 10^{43}, giving a total entropy of 10^{123}. That gives us a far larger region for the allowed phase space. If the Creator had aimed at random within this allowed phase space we'd have just seen a mass of congealed black holes (or white holes, their time-reverses). This would not be remotely like the universe we see.

At this point, some people would appeal to the so-called *anthropic principle*. That means: we're here. People are around. They exist and need a decently low level of entropy. So that restricts the possibilities. You've got to think about universes in which people exist. But the trouble with that argument is that you don't need much of the universe to exist in a low entropy state in order to get us. You can get away with just a small pocket of low entropy. The cheapest way of making a roomful of people, or even a world full of people, is by random selection. You start with particles sort of coming together at

random — and perhaps ten minutes ago that's what you were — and now suddenly you're reading this book. That's the most likely way, except that it would most probably not be ten minutes but half a minute or even less. The most likely way in which we came into existence is by some such procedure. By a statistical fluctuation, if you like, in an otherwise high entropy universe.

Well I don't believe a word of it; and I hope you don't either. Consistently with our being here, the entropy could be vastly higher than it actually is. If that's the only argument you're using, if you're arguing that the fact that we're here is why the entropy is as low as it is, then the answer to the argument is that you don't need to make the rest of the universe to make us. Given that we're here, the most probable thing outside is something completely different from what we see: namely, lots of massive black holes. When you looked outside your room, you wouldn't expect to see a lot of stars; you'd expect to see black. The anthropic argument won't work here.

So in fact, the Creator must have been constrained in building this universe. How much is the Creator constrained? Well, at least within the small region shown on the fig. 4.5. As I remarked earlier, I've drawn it much too big. By how much have I drawn it too big? Well, bearing in mind that there's a logarithm involved in entropy, it turns out that the ratio of the volume of whole phase space to the small region is $10^{10^{123}}$. You'd have to aim pretty close to hit a point that small. Now I'm not saying that the Creator wasn't that clever. What I'm driving at is that we need a *law* which says that the universe had to be like that. And why not, because, after all, aren't singularities places where ordinary physics gives up? We don't know what's going on in singularities. It may well be that the physics which applies to singularities incorporates such a principle. If so, big bangs are constrained in this kind of a way.

Here I want to postulate something I call the *Weyl curvature hypothesis*. It's no more than a hypothesis, as I've just said, but it works pretty well, I think. My point of view is: all right, you've got singularities in the universe, places where we don't understand the physics because the curvature has gone to infinity. But there's *something* physical that's going on. There must be, because of this very very strong constraint upon the entropy state. We don't understand that physics. But it could well be that that physics is such that the curvature is of a particular kind. Now curvature is described by a thing called the Reimann tensor, and relativists like to split this into two pieces: the Weyl tensor and the Ricci tensor.

The Reimann curvature splits up, naturally, in this form. Now if we take black holes, for instance, what happens, near the end, to the curvature? Well you find that the Ricci part is completely dominated by the Weyl part. The Weyl part is much much bigger than the Ricci part. This is because of distortions that are involved, leading to complications at the singularity. But according to the standard big-bang models (the so-called Friedmann models), quite the opposite is the case in the initial singularities. The Weyl part is much much less than the Ricci part. Of course you can't see these singularities. That's what one deduces, if you like. What one sees, in fact, is the expansion and the black-body radiation, the latter being a relic of the hot big bang.

We don't yet understand the physics of singularities. It's probably a part of *quantum gravity*. But I'm hypothesizing that this will turn out to have, as one of its implications, that in *initial*-type singularities the Weyl curvature is zero. That's a constraint which I'm suggesting this physics imposes. Now this is a time-asymmetric law. I said that the laws of physics that we know and love are time-symmetrical. Well, we don't understand and can't love all the laws of physics, because we don't know them. But these unknown laws, whatever they are, should, I claim, have this implication that the Weyl curvature should be zero, or close to it, in initial singularities.

We're singling out this distinction between initial and final; that has to do with which side of the singularity the part of the light cone that we call the future cone, or the *plus cone*, lies. The plus and minus cones are only different if one believes that there is a time-asymmetry behind them somewhere, in the physics. Well, we know there is. When I say we know there is, it's not utterly and completely clear; but this seems to be a strong implication of the decay of K° particles, which exhibit time-asymmetries in their behaviour. And this appears to be a fact. Of course this may not have anything to do with what I'm saying. It's just that nature seems to have this peculiar property that most of the laws that we see around us are time-symmetric. Yet they're not all *quite* time-symmetric. And the macroscopic behaviour of things seems to be grossly time-asymmetric. I'm saying that if one postulates this Weyl curvature hypothesis, just provisionally, then some of these peculiarities can be explained. The hypothesis tells us, among other things, *how* the entropy was low in the big bang. This is because the Weyl curvature measures the gravitational degrees of freedom. That is basically what it's doing. The Weyl curvature hypothesis says that, for some reason, the

gravitational part of the entropy was not at all at its maximum at the time of the big bang but was set, essentially, to zero. Somehow the gravitational part of the entropy, the space–time structure part, was not thermalized. Because gravity couples so weakly with the rest of matter, you get away with that for an awfully long time. And only very, very gradually does the gravitational entropy start to build up.

In fact, you might ask: why are we all here, why are all the things around us here? Well, one reason we're all here is that the Sun's out there, and that's a low entropy region. We receive from the Sun a small number of high-energy photons which we busily convert into large numbers of low-energy photons. Correspondingly, the entropy goes up, because there are many more degrees of freedom in the greater number of low-energy photons than in the smaller number of high-energy ones. That's what plants do for us: they sit there and take up these high-energy photons and spew out many more lower-energy photons. That's the way they grow. Then we eat them or we eat animals which eat them and so on. That's how we get around and exist without succumbing to the second law of thermodynamics.

But that raises a further question: why is the Sun there? The Sun is there because it has condensed out of a gas. Previously we had a uniform distribution of matter. If the distribution hadn't previously been uniform we couldn't have ended up with a universe such as we actually see around us. The big bang itself was, in fact, very uniform in structure, as I remarked earlier. Space itself was uniform, with an initially even distribution of gas. And it took a long time before that gas started to clump. It couldn't begin to do so until the expansion had started to die off. Then, as the expansion slowed, you began to get clumping into galaxies and stars and so on. So that all makes sense. But there is no such constraint at the end. The entropy just goes sailing up, and the entropy becomes absolutely huge. You have a very complicated singularity structure. It is nothing like what happens at the beginning.

Well, as I said earlier, some people might consider what I have been saying to be controversial. I'm not sure I understand why. I consider what I've been saying so far to be the uncontroversial part of my thesis. I want to proceed now to the part I admit to being controversial. Where might the Weyl curvature hypothesis come from? I don't think we should just pull it out of a hat, though that's one way of doing it. So let's see if we can do any better. Well, it seems to me that the Weyl curvature hypothesis must have something to do with the, at the moment, non-existent theory of quantum gravity.

That's a rather vacuous statement. But a lot of people are striving towards attaining a theory of quantum gravity, and fine: good idea. But the way in which people try to do this, on the whole, is to take gravity, general relativity, and take quantum mechanics. First of all, they try and do a sort of quantizing of the classical theory, which means applying the rules of quantum mechanics to general relativity and seeing what they get out. Well, I should say that these attempts have been rather dismal in their results. There's a lot of work, a lot of very good work on the subject, but no viable theories seem to come out of it. In fact, the theories that emerge appear to tell you that there's something wrong, that straight general relativity combined with quantum mechanics leads you to something which is *non-renormalizable*, in which totally intractable infinities appear. (Infinities appear in standard quantum-field theory, too. But here there is an *ad hoc* procedure, known as *renormalization*, for dealing with them. Non-renormalizable theories are ones that generate infinities which are not susceptible to being 'tamed' in this way.)

This leads some people to say: well, perhaps we should change general relativity. And perhaps we should. But my own prejudices are the other way around, that if we have this kind of conflict, we ought to do something about quantum mechanics. I'm not saying it's completely the other way around. No doubt one will have to do *something* about general relativity. It certainly can't be totally correct, as it stands. After all, you can't have a classical theory and a quantum theory just sitting there together like that. They have somehow to interact with each other. So general relativity will have to give a little: it will have to have a quantum form. But I believe that quantum mechanics likewise will have to give a little, in fact probably give quite a lot. And one of the things I have against quantum mechanics, and that a lot of people have against quantum mechanics, is that it doesn't make any sense.

Perhaps I should qualify this. It is, of course, an absolutely marvellous theory. It's marvellous for at least two reasons. One reason is that it agrees with all experiments, enabling you to do all sorts of things that couldn't be done before, and you get marvellous explanations and predictions of phenomena coming out. I'm in no position to talk about many of these, because my understanding of them is inadequate. But the ones I do know about are pretty impressive. There's obviously a lot that's right about quantum mechanics in observational terms. The other thing that's marvellous about quantum mechanics is its mathematical elegance. Mathematic-

ally, it is a very elegant theory. I should say there's no question about that. So I would certainly not wish to see quantum mechanics replaced, just like that, by something else. And I don't believe that it will be. I think that it has marvellous strengths, both from the theoretical and the experimental side.

Of course when someone starts with this sort of eulogy, you can tell that he wants to be rude about the subject! Well I've already been rude: I've said that quantum mechanics doesn't make sense. But then you see it *doesn't* make any sense. It's all right when you're looking at simple things like single particles or simple systems of particles. But, as I think most of you know, there are things like the Schrödinger's cat paradox and so on which seem to indicate that, when you have a large structure classical physics applies, and something other than the linear rules of quantum mechanics seems to operate. Is the cat in Schrödinger's box, in his thought experiment, really in some linear combination of dead and alive states? Or is it, in reality, *either* dead *or* alive? Well Schrödinger himself certainly thought that it was either dead or alive. And it's not just, as you often read in popular accounts, a question of its being half dead and half alive. According to quantum mechanics, it's in a *complex* linear combination of dead and alive. It could be dead plus alive. It could be dead minus alive. It could be dead plus the square root of minus one times alive. All these things are different and, in principle, distinguishable physically. Now that's what quantum mechanics appears to say and it's difficult to take seriously; because that's probably not what happens. But if it's not what happens physically, then there has, at some stage, to be a modification in quantum mechanics.

The question, then, is what form this modification should take. Before giving you my ideas about this, let me give a more precise statement of the Weyl curvature hypothesis. It seems to be possible to divide all singularities into two mutually exclusive and jointly exhaustive classes: past singularities (such as the big bang) and future singularities (such as the big crunch or the singularities contained in black holes). What the hypothesis then says is: *the Weyl curvature tends to zero at all past singularities, as the singularity is approached from future directions*. Two things not so far mentioned follow from this. First, the principle rules out the existence of so-called *white holes*. These are the hypothetical time-reversed counterparts of black holes: a white hole would spew out matter just as black holes suck matter in. Just as you can't get out of

a black hole, so it would be impossible, without exceeding the local light velocity, to get into a white hole. There have been a number of intriguing speculations concerning white holes. One is that we are actually living in a white hole. That is observationally wrong, since we have seen that the big bang singularity is actually not at all like that in a black hole. Another speculation is that the matter sucked in by a black hole at one point in the universe is emitted by a white hole at some other point, perhaps thousands of light years distant. Well, I'm sorry to be a spoilsport. But the Weyl curvature hypothesis unambiguously rules this out too.

The second and related thing that follows is that we get a problem when we consider regions of phase space that are supposed to contain black holes. Once we rule out white holes, it turns out that there are more paths going into such a region of phase space than there are going out. And this violates something known as *Liouville's theorem*. This is because lots of widely divergent conditions at one time can, if they lead to gravitational collapse into a black-hole state, result in identical states at a later time. Black holes, as the saying goes, 'have no hair': they retain no trace of their precise origins. What I want to suggest is that a proper quantum theory of gravity, once we have such a thing, will allow for actual physical collapse of the wave function; and that this physical collapse compensates for what happens elsewhere in the phase space, when a black hole forms. Just as there are more paths into than out of a point in phase space corresponding to a black-hole state, so there would be more paths out of than into a point of phase space corresponding to a physical collapse of the wave function. When we have both phenomena, this allows for an overall balance within the total phase space. For the balance to be achieved, we require that wave function collapse be a *gravitational* phenomenon. An important point is that quantum mechanics tells us that identical wave functions may collapse into quite different states. For example, in the case of Schrödinger's cat, half-dead plus half-alive may collapse either into dead or into alive. Indeed, the whole point of what I'm now saying is that it *must* do so, in this sort of case. Big systems cannot, if I am right, exist in linear combinations — *superpositions* as they are called — of macroscopic states.

It remains to consider under what conditions, precisely, the wave function should be expected to collapse. Well, from a certain point of view, a physical collapse of the wave function ought to be associated with a fall in entropy — if we ignore gravity, that is. (This is a

'commonsense' entropy and not 'quantum-mechanical entropy'.) That would explain why we don't get spontaneous collapse in very small systems: for them gravitation is of negligible strength. But as we consider progressively larger systems, gravitation becomes progressively more significant. Moreover, gravitational entropy, as we have seen, rises as matter becomes more concentrated. So it's a question of balancing these two effects. My suggestion is that the wave function collapses — or at any rate is physically *allowed* to collapse — when the system is sufficiently massive for such collapse to result in a net gain of entropy. In other words, the wave function only collapses when it can do so consistently with entropy going up, or at least not falling. To calculate when this would happen, one needs a general way of measuring gravitational entropy. The Hawking–Bekenstein formula is just one example of a formula that assigns an entropy to a space–time geometry; but we want to be able to do this quite generally. There is a particular formula that I have been working with, which is a generalization of the Hawking–Bekenstein formula; and I have tried applying this to the problem of droplets condensing in a cloud chamber. What I end up with is figures around a millimetre for the diameter at which a physical collapse of the wave function should be expected to occur. This is perhaps a bit on the large side, but not ridiculously so. (It would suggest that anything from fleas on down can exist in linear superpositions of states, but cats certainly can't.)

I must emphasize that these ideas are highly speculative and not at all fully worked out. But I find them suggestive. There is something deeply unsatisfactory about the present conventional formulation of quantum mechanics, incorporating as it does two quite distinct modes of evolution: one completely deterministic, in accordance with the Schrödinger equation, the other a probabilistic collapse. And it is a great weakness of the conventional theory that one is not told when one form of evolution is supposed to give way to the other, beyond the fact that it must always take place sometime prior to an observation being made. (Some, like Wigner, have even suggested that it is consciousness that collapses the wave function.) The foregoing line of thought holds out some promise of our being able to synthesize these two elements into a single, unified theory. Obviously, if I am right, the Schrödinger equation will have to be modified in some way. And just how I am not sure. It is even possible that we may end up restoring determinism in quantum mechanics, which would have pleased Einstein. Personally though, I

doubt that it would be quite like that; and it is not, in any case, the statistical character of contemporary quantum mechanics that creates the problems. What we are after, as I say, is a theory that (a) accommodates gravity and (b) makes sense. Some people seem content to regard quantum mechanics simply as a useful computational tool, not worrying about its interpretation. But, as a philosophical realist, that doesn't satisfy me. I cannot believe that it is impossible to make sense of the physical world. I think that it is the task of physics to do precisely that.

POSTSCRIPT

A while after giving the lecture on which this article was based, I somewhat shifted my view as to the criterion determining physical wave-function collapse (partly as a result of objections by R. M. Wald). I do not now think that the criterion should be phrased in terms of a concept of 'gravitational entropy' but, rather, the more clear-cut 'graviton number'. On my *present* view, wave-function collapse takes place when the difference between the gravitational fields of the states under superposition amounts to at least one graviton's worth. Combining this reinterpretation with a refined calculation by A. Ashtekar, we can reduce the mass of the droplet by a factor of about a million — a considerable improvement.

FURTHER READING

Landsberg, P. T., *The Enigma of Time* (Adam Hilger, Bristol, 1982). This book contains an excerpt which is reasonably accessible to the non-specialist from R. Penrose's paper 'Singularities and time asymmetry' which appears in full in: Hawking, S. W. and Israel, W. (eds) *General Relativity — An Einstein Centenary Survey* (Cambridge University Press, Cambridge, 1979).

Layzer, David, 'The asymmetry of time', *Scientific American*, Dec. 1975. This article discusses further the conflict raised by Dr Penrose between thermal equilibrium and increasing entropy and defends the viewpoint that Dr Penrose questions.

Further details of the Hawking—Bekenstein formula may be found in the following technical articles: Hawking, S. W. *Communications in*

Mathematics and Physics, 43 (1975) p. 199 and Bekenstein, J. D., *Physical Review*, D7 (1973), p. 2333; 9 (1974), 3292.

A proposal to introduce consciousness directly into quantum mechanics is given by Eugene Wigner in *The Scientist Speculates*, ed. I. J. Good (Heinemann, London, 1962).

168
ESCHER AND THE VISUAL REPRESENTATION OF MATHEMATICAL IDEAS

With a note by M. Emmer. *M. C. Escher: Art and Science – International Congress Proceedings* (1985), 143–57, North-Holland, Amsterdam, 1986

In part, this is a historical account of my encounters with the remarkable Dutch artist M. C. Escher. It also shows my depiction of the paradoxical issues that arise when certain kinds of 'impossible object' are drawn with full perspective (Figure 7, on p. [152]). Compare this with Figure 3 on p. [147] of Chapter 10. At this conference I became involved in a conversation with an American mathematician participant, to whom I explained that these pictures were illustrations of *cohomology* (see Chapter 229). He asked whether one could illustrate this with other groups than the $\mathbb{R}^+$ that is employed here. I thought about it, and the next day, at my talk, produced an example for which the group is $\mathbb{Z}_2$. Other examples using $\mathbb{Z}_2$ appear here: Figures 9 and 10 on pp. [155] and [156].

526 Escher and the visual representation of mathematical ideas. With a note by M. Emmer.
M. C. Escher: Art and Science – International Congress Proceedings (1985), 143–57, 1986

ESCHER AND THE VISUAL REPRESENTATION OF MATHEMATICAL IDEAS

Roger Penrose

Mathematical Institute
University of Oxford
Oxford OX1 3LB
United Kingdom

In the art of M.C.Escher can be found many marvellous illustrations of abstract mathematical ideas. The providing of such visual realizations of mathematical concepts can be a valuable aid to mathematical understanding and even to future research. Symmetry groups, periodic and non-periodic tilings of the Euclidean and Lobachevski planes,self-referential logical concepts and paradoxes,impossible (but locally possible) objects with their relations to global (i.e. holistic) mathematical concepts; all these and other mathematical ideas are vividly illustrated by Escher and by some others who have followed him.

In the work of Escher we find a fascinating interplay between visual art and numerous mathematical ideas.The aesthetic qualities of his art enrich,and are enriched by,the mathematical qualities of logic,geometry and paradox.Much study has been made of Escher's plane tessellations,where he used tiles of fascinating shapes:birds, fish,frogs and innumerable other creatures and objects to illustrate the various two-dimensional crystallographic groups - the discrete subgroups of the group of Euclidean motions of the plane.I do not propose to say anything about this here since the matter is well covered in the articles,in this volume,by several authors.I shall,however,return briefly to a somewhat different aspect of plane tilings in a moment.Escher also had some interest in three-dimensional (space-filling) tessellations ('Flatworms') and in the two-dimensional tessellations for which the Euclidean plane is replaced by the surface of a sphere('Sphere with Men','Sphere with Angels and Devils') or by the hyperbolic(Lobachevsky) plane (the various 'Circle Limits'). The latter opens up opportunities for illustrations of Riemann surfaces(of genus 2 or more) where one 'wraps up' the hyperbolic plane into a closed surface(which might have the topology of,for example,the surface of a pretzel) and the Escher design can,in a consistent way,cover the whole surface.(I have occasionally seen examples,that people have constructed,of this kind of thing).

Escher's illustrations of the hyperbolic plane serve also to give

one a feeling of a mathematical concept of infinity-as also do some of his other woodcuts('Square Limit','Whirlpools','Smaller and Smaller',etc.)-where more and more of the creatures in these designs crowd up against the edge,or interior point.Also the logical paradoxes arising in the mathematical study of the infinite are nicely illustrated in Escher's work through the self-referential aspects of two of his prints('Fishes and Scales','Print Gallery'). Escher was fascinated by polyhedra-not just the original Platonic (convex regular)five,but also others,including non-convex(Kepler-Poinsot)examples.He was interested in,and made use of,many other geometrical constructions,being continually intrigued by such things as the appearance of objects reflected in spheres;and he touched upon much else of a mathematical nature.It is hardly surprising that his work is so beloved of mathematicians,and of the mathematically inclined,throughout the world.

I should like to say a little more about tilings of the plane.It has been pointed out [3] that in Escher's final black and white and orange water-colour,he has depicted an example of a non-isohedral tiling of the plane.Thus,the single shape-out of which the entire pattern is constructed and which covers the whole plane without gaps or overlaps-is placed in more than one distinct way in relation to the pattern as a whole.That is to say,not all occurences of the shape are on an equal footing:one can find two instances of it such that there is no Euclidean symmetry,taking the pattern into itself,which carries the one instance of the shape into the other.Moreover,there is no alternative way of tiling the plane with this particular shape.It will tile only non-isohedrally.

This appears to be Escher's only example of a non-isohedral tiling. Actually,he got the idea of this tiling(although not the particular shape of piece)from me.My wife and I had paid an unexpected visit to Escher at his home in Baarn in,I believe,1962.He was most friendly to us and he let me choose one of his prints.I left him behind some wooden puzzle pieces that I had made-all identical in shape.The shape is illustrated in Figure 1b.The puzzle was to find a way of tiling the entire plane with this shape.Some short while later Escher wrote to me enclosing a diagram with the solution of the puzzle,but expressing some surprise that the solution might be,as I had asserted,unique.

Now I have always been a poor-or,at best,erratic-correspondent.It must have been well over a year before I replied.It may even have been a good deal over two years.I still feel guilty about this delay.When I finally did respond,I explained the principle on which the design of the wooden pieces had been based.(Fig.1a,Fig.2) I also illustrated this principle with a bird-like shape which would tile the plane only according to this scheme(and therefore only non-isohedrally).Some while later Escher wrote back in reply,explaining that he had not been well,but enclosing a copy of the shape(Fig.1c)that now appears in his final(non-isohedral)picture.He was rather apologetic about it and referred to it as his 'little ghost'.*

* See page 395 (top left) for colour reproduction ('Ghosts').

 Escher and the visual representation of mathematical ideas. With a note by M. Emmer. *M. C. Escher: Art and Science – International Congress Proceedings* (1985), 143–57, 1986

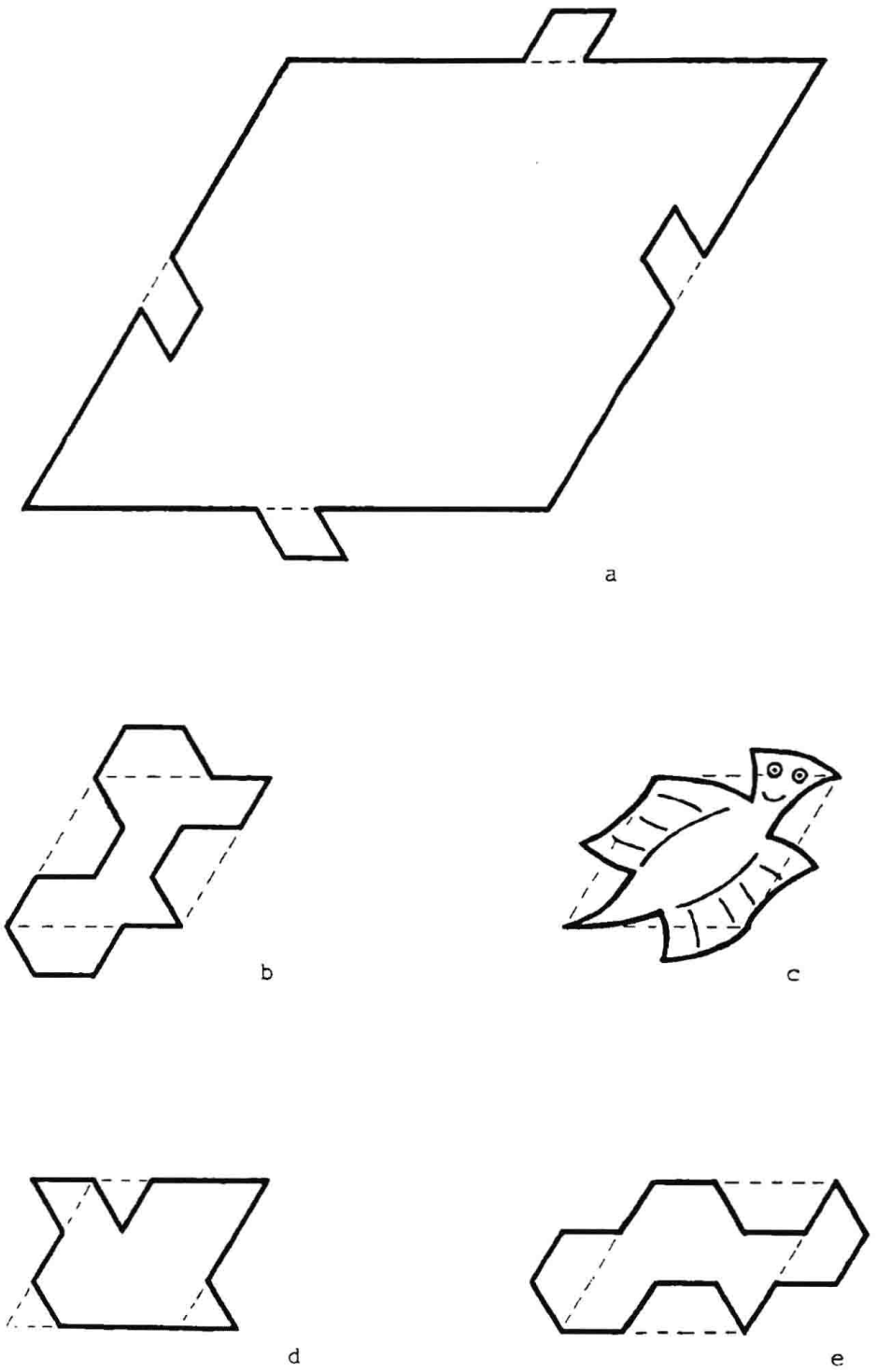


Figure 1

I had already published [9] ,with my father,an example of a shape which would tile the plane only according to this scheme(Fig.1d),and other examples of shapes which would tile the plane only non-isohedrally appeared in that article.(Other examples had been known for some time)The puzzle pieces that I had left with Escher were designed to have many 'false tries' so that the correct assembly pattern of the pieces would not be immediately obvious.At that time I had also known of another similar shape with even more 'false tries',but I was not yet certain of the uniqueness of the tiling for that shape(Fig.1e).Some while later I convinced myself that the tiling pattern was indeed unique;and,in 1975,Martin Gardner set it as a problem in his Mathematical Games column in the Scientific American [1] ,referring to the shape as a 'loaded wheelbarrow'.

In another of his column [2] ,Gardner described a tiling of mine that is not even periodic,let alone isohedral.Here two distinct shapes are needed,either the kite/dart pair or the rhombus pair illustrated in Figure 3.An assembled region,using each pair,is shown in Figure 4.

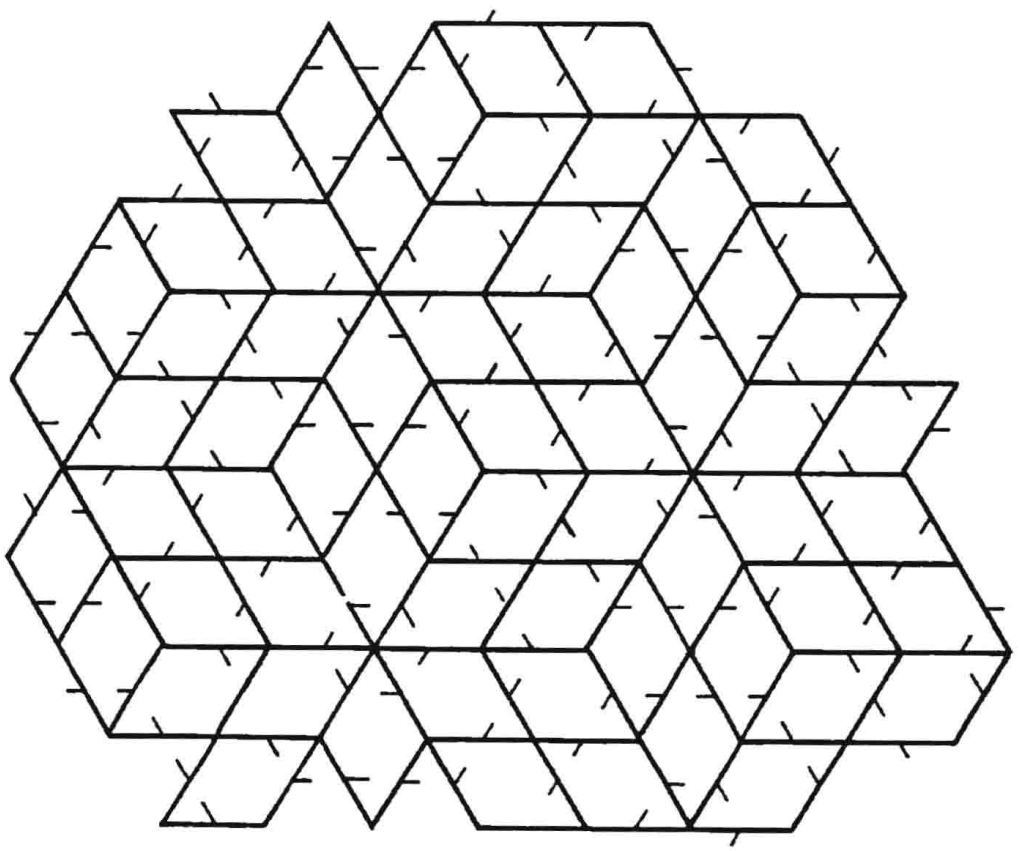


Figure 2

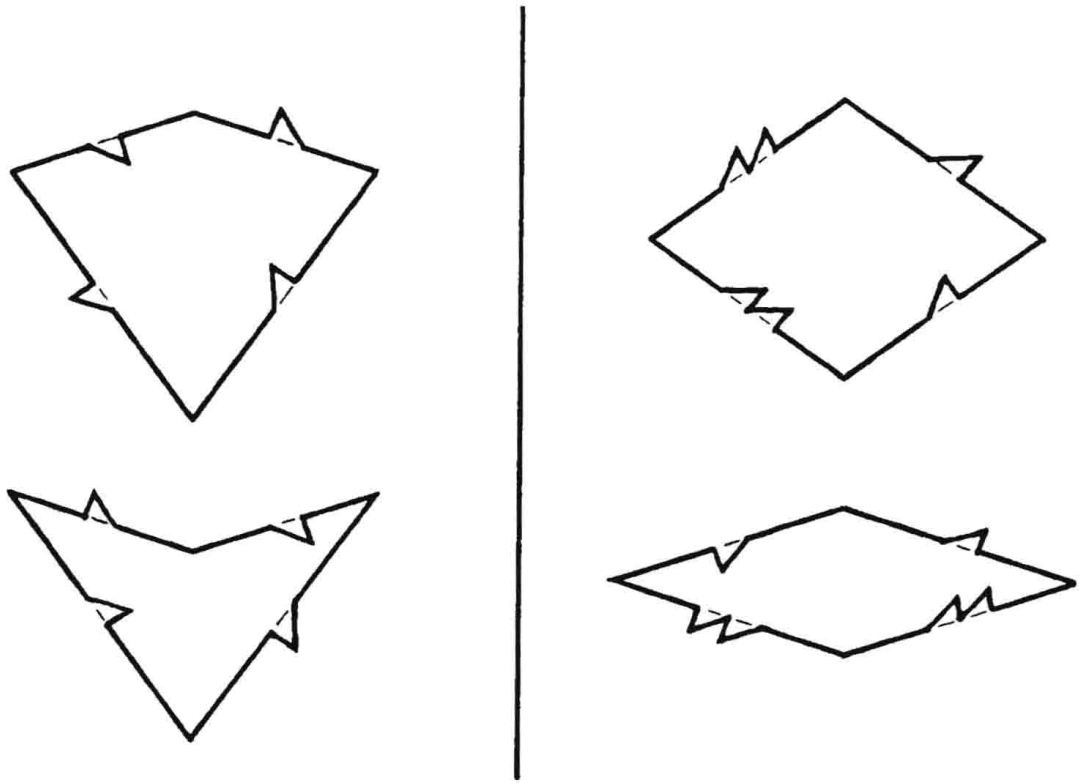


Figure 3

Like the single tiles that I have been describing above,one can modify the edges of these pairs in many ways to provide a variety of interesting shapes,such as the pair of birds illustrated in Figure 5.(1) It is one of my very great regrets that Escher did not live to make use of these tiles.One cannot adequately speculate on the fascinating designs that he might have produced.(Moreover,I have a personal regret that my father,also,did not live to see them.He died at just about the same time and at about the same age as Escher) I have described my non-periodic tiles(and their genesis)elsewhere [10] .They were based on an earlier six-piece scheme [11] .I do not propose to say more about them here.They,and other non-periodic tilings are extensively discussed in a new book by Grünbaum and Shephard[4].

Within the last year a startling discovery has been announced. Schechtman and his collaborators[13]appear to have seen,in certain alluminium-manganese alloys,that a characteristic five-fold(or ten-fold)symmetric arrangement arises in the electron diffraction pattern,which is impossible for any conventional(periodic)crystal.

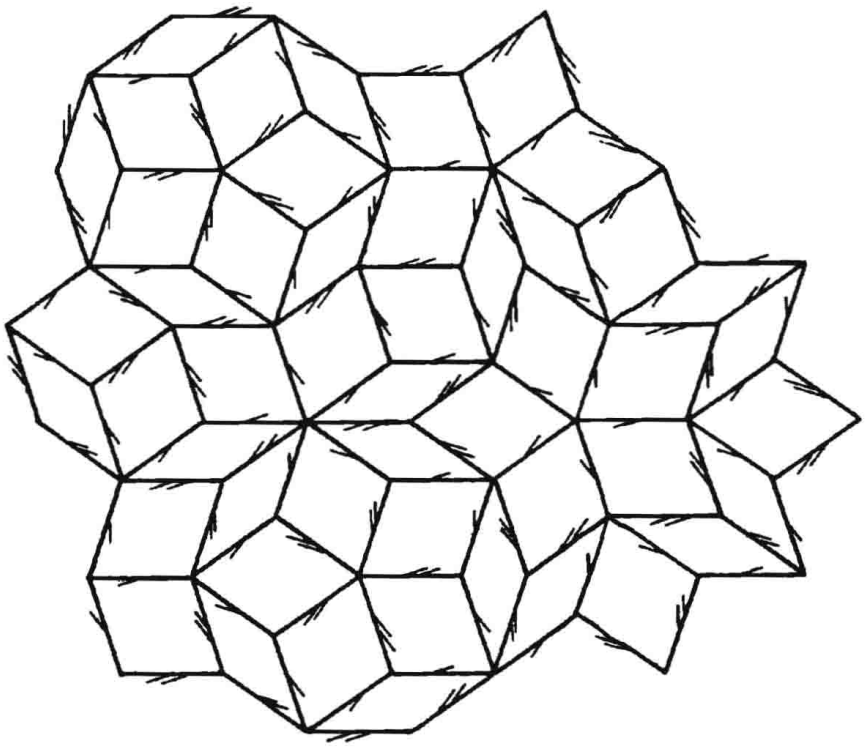


Figure 4

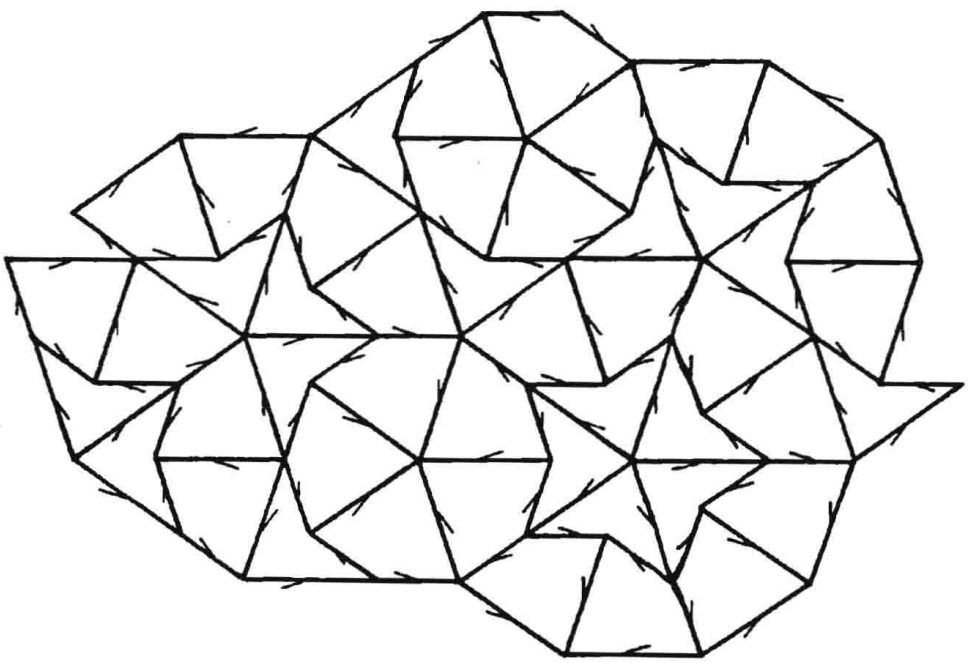


Figure 5

As Levine and Steinhardt [5] independently pointed out,precisely such arrangements would indeed arise if the atoms were organized according to one of my non-periodic patterns-or,rather to the natural three-dimensional generalizations of such patterns.Somewhat earlier MacKay [6] had made similar,though somewhat less clear-cut suggestions,using a three-dimensional version of my tilings described,quite early on,by Robert Ammann.Various popular accounts of these discoveries have now appeared. [12], [7] It seems clear that the experimental discovery of these quasi-crystals has raised many puzzling questions and opened up a fascinating new field of enquiry.

The rest of what I want to say concerns impossible objects.My own involvement in these things dates back to 1954,when I attended the international Congress of Mathematicians in Amsterdam as a research student.A lecturer of my acquaintance suggested that I would be interested in an exhibition of work by the Dutch artist M.C.Escher,this exhibition having(apparently)been specially arranged to coincide with the Congress.I went along and,as I recall,was totally fascinated,never having come across Escher's work before.On returning to England,I decided to try my hand at something impossible myself.I tried various designs with rods passing behind and in front of each other in various ways.Finally I came up with the impossible triangle(i.e.,with what has in more recent years become known as the tribar)which,in my view,embodied the impossibility I was trying to express,in its purest form.Although,in his exhibition,Escher had had many strange and wonderful things,there was nothing that we would now call an impossible object quite in that sense.

When I next saw my father,I showed him the triangle.He at once developed it in various ways-I recall,particularly,his impossible college-until he produced a picture of an impossible staircase which went down(or,equivalently,up)all the way around.We felt that we would like to publish these things,but we had no clear idea how to do this,not really knowing what the subject was.Then my father recalled that he happened to know the editor of the British Journal of Psychology and could probably get him to accept our article.So we decided that the subject must of course be psychology,and we sent the paper there.Eventually it was published in 1958.[8] We had given acknowledgement to the catalogue of the Escher exhibition that I had attended in Amsterdam;and when offprints arrived we sent Escher one,expressing our appreciation.

In the meantime,and quite independently,Escher had produced his lithograph 'Belvedere' which illustrates a very similar kind of impossibility.Some while later recieving our offprint,Escher produced 'Ascending and Descending',and then,rather later,'Waterfall'.Escher was full in his acknowledgement of our paper.We where none of us aware of the much earlier work of Oscar Reutersvärd.Indeed I only came across it myself some two years ago.(See the article by Ernst in this volume for an account of early impossible objects).

There was one feature that our tribar picture posessed which I have not seen in Reutersvärd's work,nor in many other impossible drawings,and that is perspective.I have given a rather more exaggerated example in Figure 6.My father and I were fascinated by some of the absurd implications of this.As one works ones way around the picture,owing to the inconsistency in the distance away that the structure is perceived to be,one finds an inconsistency in the size of any object associated with it.In our 'Puzzles for Christmas' article[9]we presented a(deliberately)rather incomprehensible puzzle,concerning the parking of a van,which illustrated this point.My faher constructed a model with some dogs going around an impossible staircase,in which the dogs incomprehensibly jumped in size(in the photograph).Somewhat earlier I had made a drawing involving a train,a ladder and a bridge,which emphasized these paradoxes of size which arise with perspective impossible objects.I drew a much improved version of this picture a few years ago in a notebook of mine.This is given in Figure 7,where I have drawn it out again rather more carefully.

I should like to point out a distinction between various types of impossible pictures which,as far as I am aware,has not been emphasized before.The impossible pictures that I have presented have all possessed a quality which,for want of a better word,I shall call pure.A pure impossible picture is made up from parts each of which has an unambiguous consistent interpretation-unambiguous,that is,apart from the inevitable(universal)ambiguity possessed by all pictures,which is the uncertainty about how far away

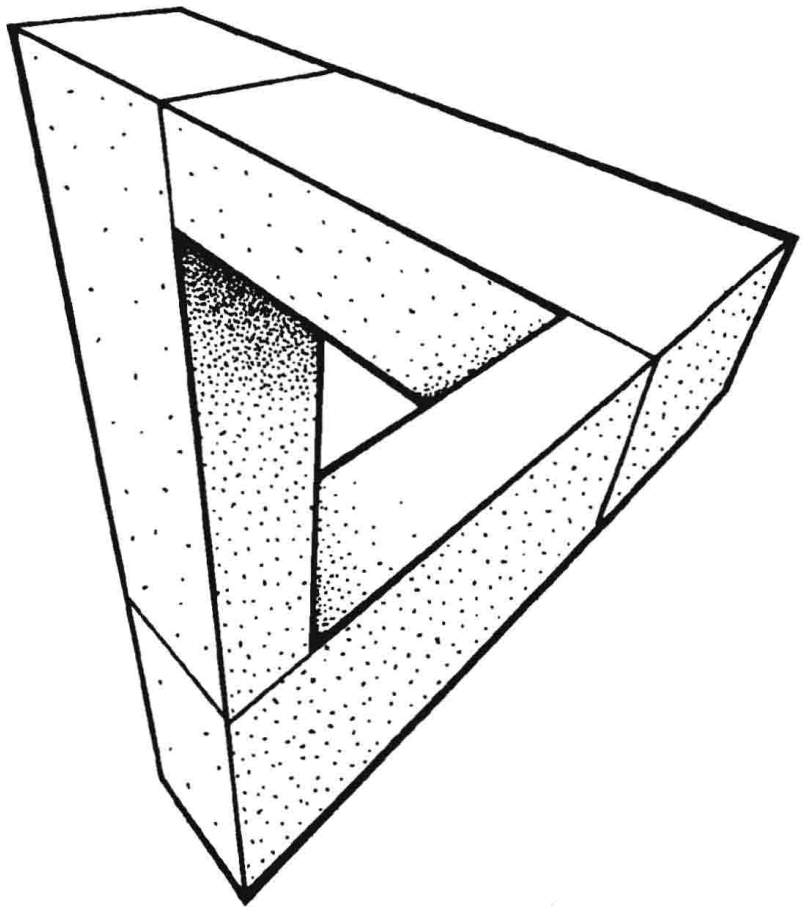


Figure 6

from the viewer the object being depicted really is.However,many familiar impossible pictures are not pure in this sense.Various examples of such impure pictures are given in Figure 8.The impurity is sometimes a little subtle,perhaps just involving the ambiguous directions in which certain sloping faces are to be interpretated.A good test for deciding whether or not an impossible picture is pure is to supply it with shading and indicate the direction of the grain in the wood,etc..If the picture is pure then as much local details can be added as desired(excluding shadows from distant parts of the structure-that is non-local)and the effect will only be enhanced.But if it is impure,there will be an inconsistency and the effect will rapidly be destroyed.

It occurred to me recently that(pure) impossible objects illustrate an important('holistic')concept in pure mathematics,namely cohomology.

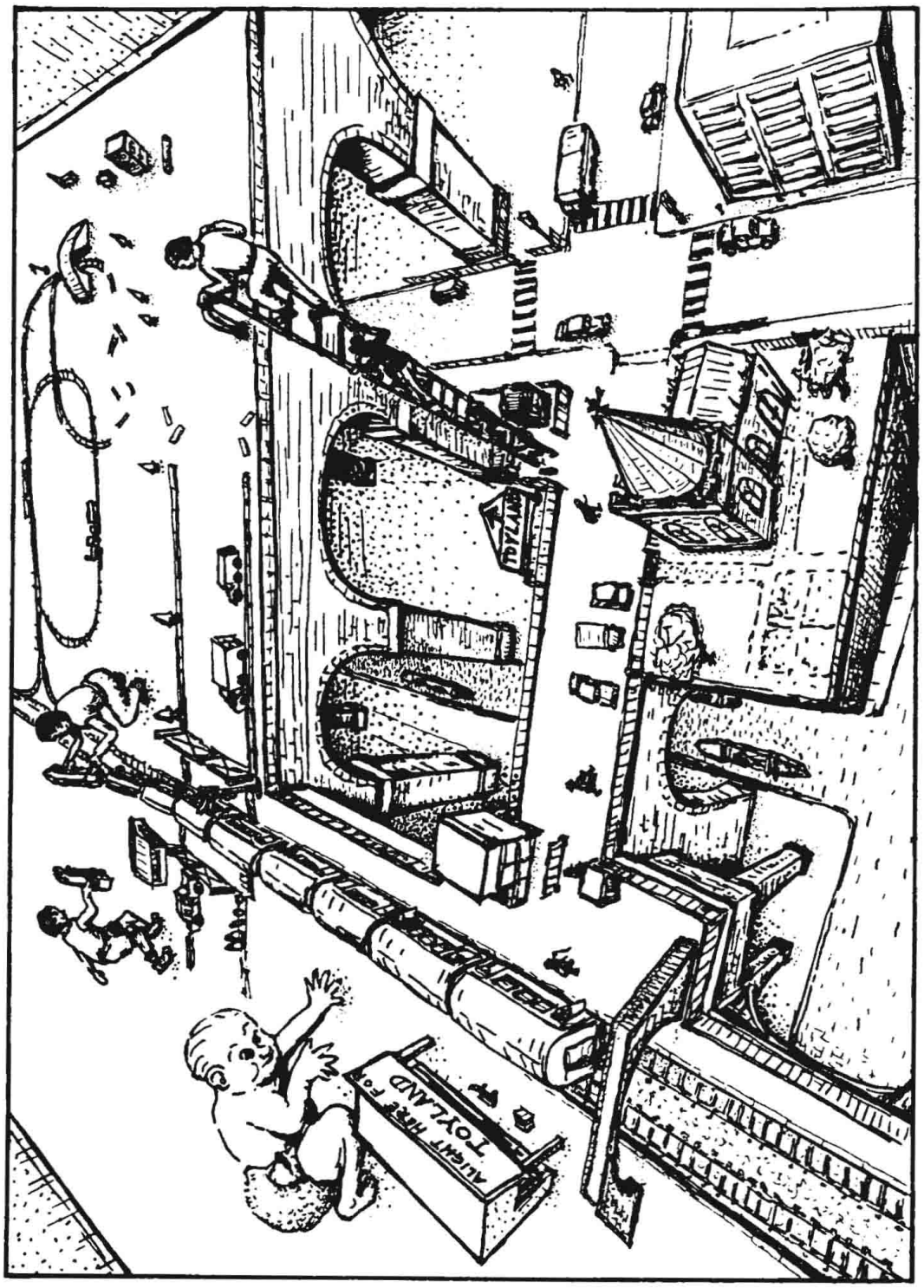


Figure 7

536 Escher and the visual representation of mathematical ideas. With a note by M. Emmer. *M. C. Escher: Art and Science – International Congress Proceedings* (1985), 143–57, 1986

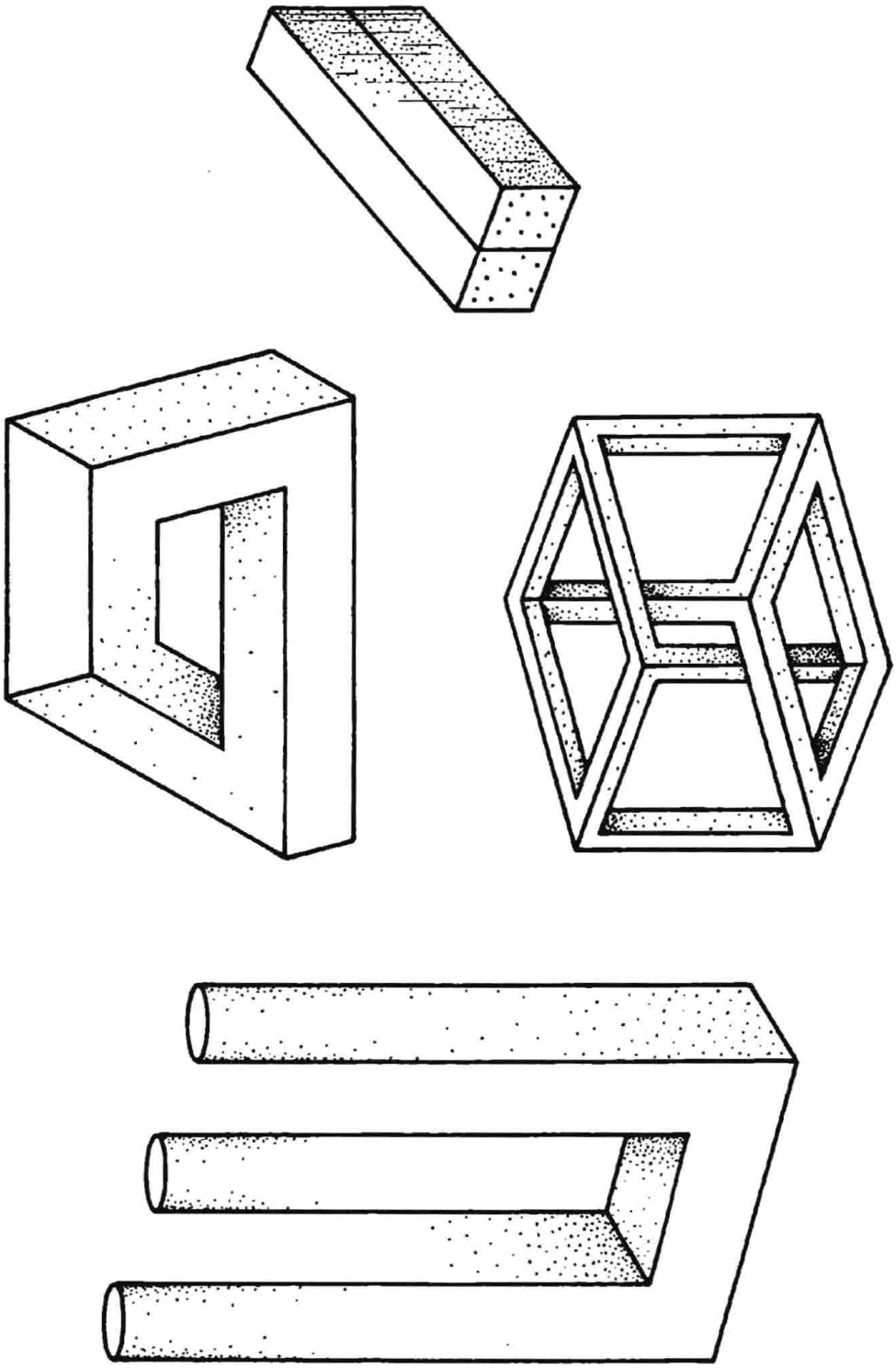


Figure 8

M. C. Escher: Art and Science – International Congress Proceedings (1985), 143–57, 1986

It would be inappropriate for me to enter into a long discussion of this here.Let me simply point out-to those who know what the words mean-that the tribar is an illustration of a non-trivial first cohomology element with coefficients in the reals(an element of $\underline{H}^1(..,\mathbb{R})$.I should like to pose as a problem that someone find a good visual illustration of a second cohomology element(element of $\underline{H}^2(..,..)$.(I have no idea whether there is one!) Stimulated by discussions at the conference it occurred to me that one can nicely illustrate elements of $\underline{H}^1(...,Z_2)$ with impossible pictures which are, in a sense,the very reverse of pure!Rather than being everywhere locally unambiguous,but globally impossible,they are everywhere locally ambiguous,yet globally impossible.Escher has something a bit like it in his highly ingenious lithograph'Concave and Convex',but that does not really illustrate the point I am getting at here.As far as I am aware,the drawings I am now presenting(Fig.9,10)represent a new type of impossible objects!I offer them as a homage to Maurits C.Escher and his wonderful work.

NOTE

(1)An animation sequence of the pattern with the pair of birds is included in the movie 'Symmetry and Tessellations'.The movie of the series'Art and Mathematics' was realized by Michele Emmer in 1980 with the cooperation of A.M.Liquori,biochemist,L. Saffaro,artist,Roman Vlad,musician,A.Vigna Taglianti,zoologist,M. Wurtz,biochemist and Roger Penrose.(Note By M.Emmer)

538 Escher and the visual representation of mathematical ideas. With a note by M. Emmer. *M. C. Escher: Art and Science – International Congress Proceedings* (1985), 143–57, 1986

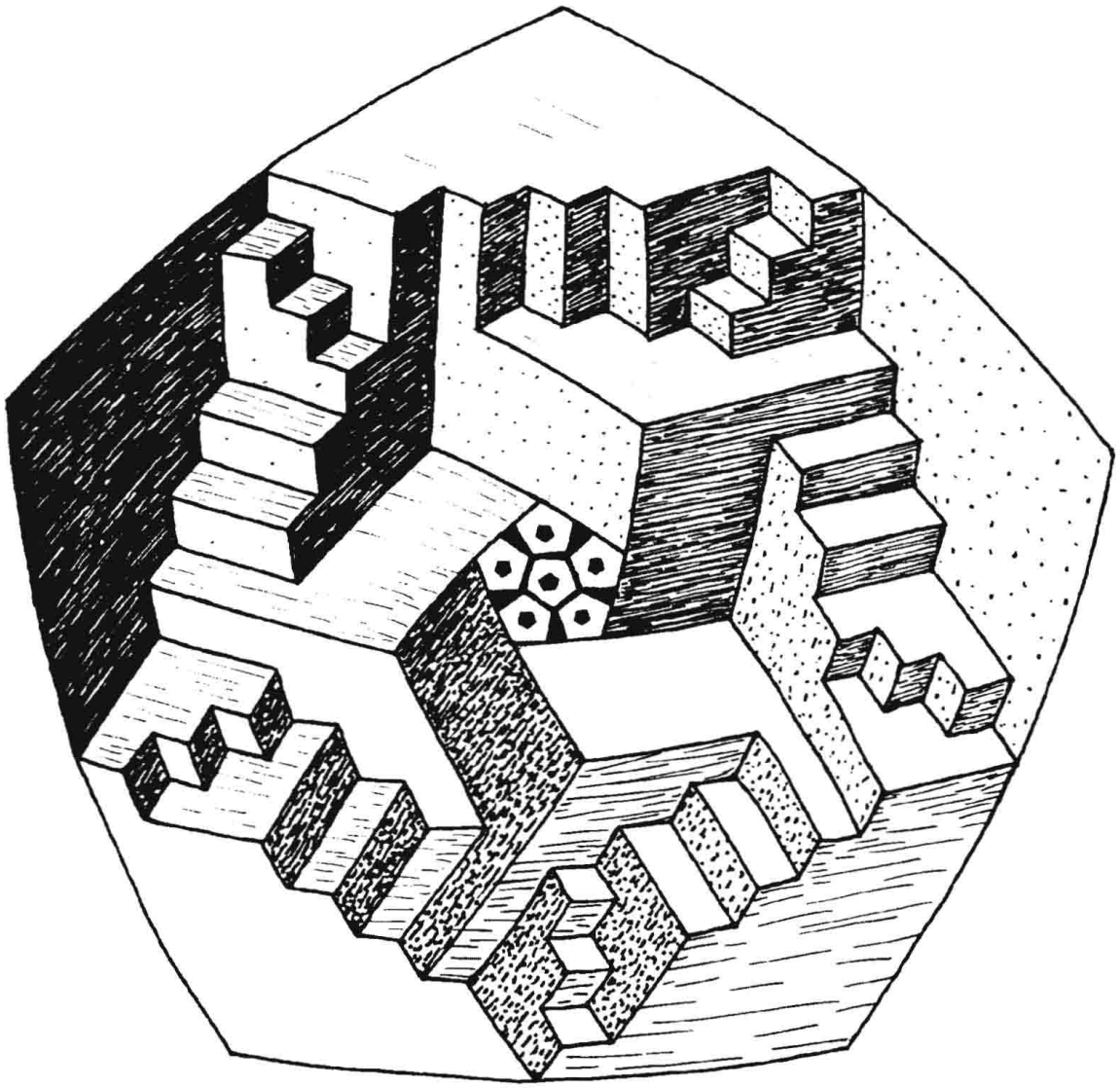


Figure 9

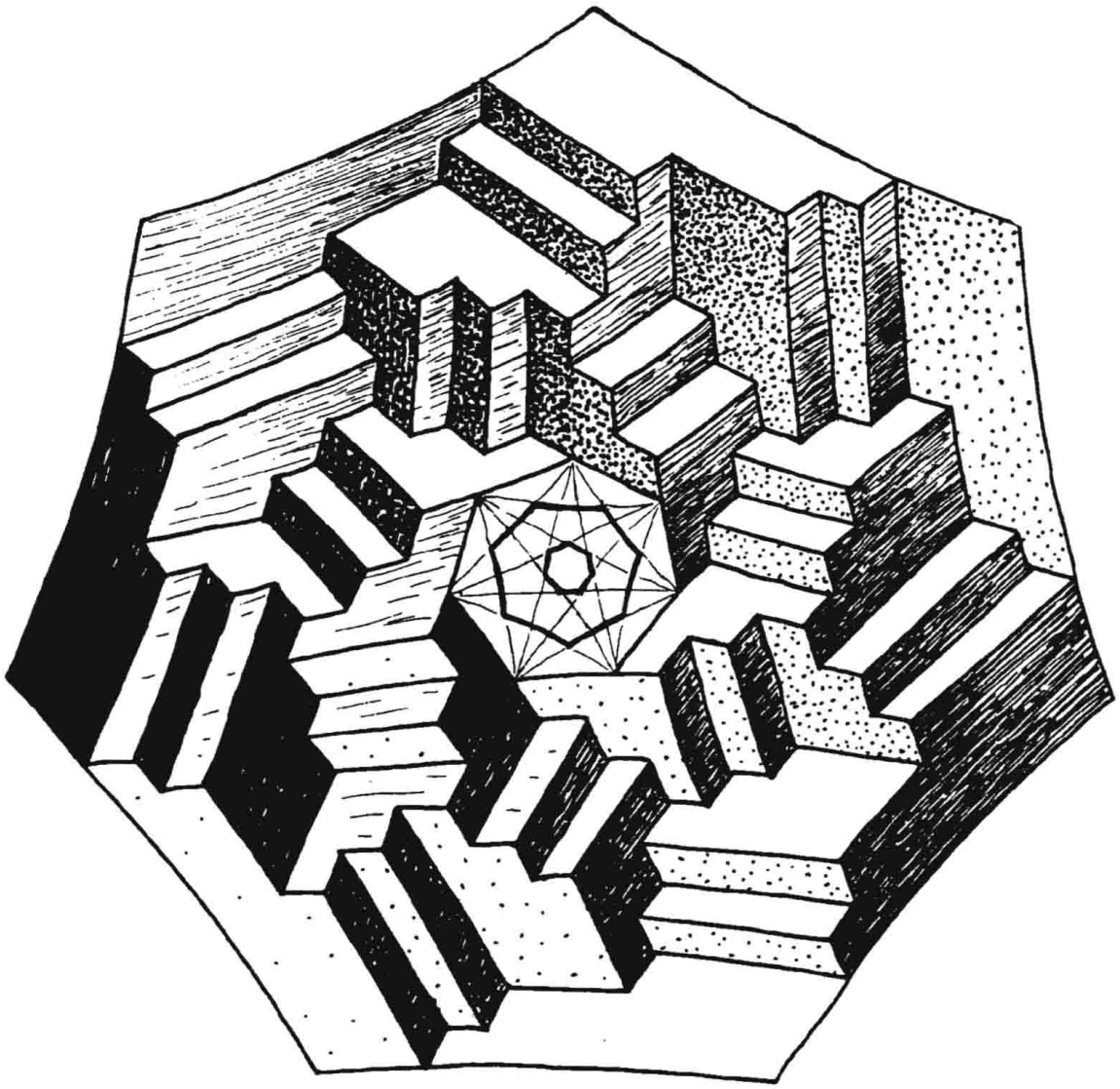


Figure 10

REFERENCES

[1] Gardner,M.,Mathematical Games,Scientific American,233(1975) 115.
[2] _____,Mathematical Games,Scientific American,236(1977)110.
[3] Grünbaum,B.,Mathematical Challenges in Escher's Geometry,in this volume.
[4] Grünbaum,B. and Shephard,G.C.,Tilings and Patterns(Freeman,S.Francisco,1985).
[5] Levine,D. and Steinhardt,P.J.,Quasi Crystals:a New Class of ordered Structures,Phys.Rev. Letters,53(1984)2477.
[6] Mackay,A.L.,Crystallography and the Penrose Pattern,Physica,114A (1982)609.
[7] Nelson,D.R. and Halperin,B.I.,Pentagonal and Icosahedral Order in Rapidity Cooled Metals,Science 229(1985)233.
[8] Penrose,L.S.and Penrose,R.,Impossible Objects:a Special Type of Visual Illusion,Brit.J.Psych.,49(1958)31.
[9] _____,Puzzles for Christmas,New Scientist (Dec.25,1958).
[10] Penrose,R.,Pentaplexity,Mathematical Intelligencer,2(1979)32.
[11] _____,On the Role of Aesthetics in Pure and Applied Mathematical Research,Bull.Inst.Math.Appl.,10,n°.7/8(1974)266.
[12] Peterson,I.,The Fivefold Way for Crystals,Science News,127(1985) 177.
[13] Shechtman,D.,Blech,I.,Gratias,D. and Cahn,J.W.,Metallic Phase with Long Range Orientational Order and no Traslational Symmetry,Phys. Rev.Letters,53(1984)1951.

169
HERMANN WEYL, SPACE-TIME AND CONFORMAL GEOMETRY

Hermann Weyl, 1885–1985. Ed. K. Chandrasekharan, 23–52. Eidgenössische Tech. Hochschule, Zürich, and Springer, Berlin, 1987

I regretted not having noticed the close association between my 'Weyl curvature hypothesis', referred to here on pp. [47] and [48] and Weyl's own 'Weyl's postulate' (see Bondi, H. (1961) *Cosmology.* Cambridge University Press, Cambridge). See also Chapter 181. On p. [34] of the present article, Figure 15 depicts my early attempt at a non-periodic 3-dimensional version of Figure 4 on p. [269] of Chapter 64.

542 Hermann Weyl, space-time and conformal geometry. *Hermann Weyl, 1885–1985.* Ed. K. Chandrasekharan, 23–52. 1987

Hermann Weyl, Space-Time and Conformal Geometry

Roger Penrose

It is a great pleasure for me to be in Zurich, with its many associations: Weyl, Einstein, Minkowski – whose work has been so close to my heart. But particularly, on this special occasion, I am honoured, privileged, and indeed very flattered, to have this opportunity to pay my respects to Hermann Weyl, whom I take to be the greatest mathematician of this century. In saying this, I am avoiding any arguments about Hilbert, Poincaré, or even Cartan, by insisting that when I say "of this century", I mean whose work lies *entirely* within the nineteen-hundreds. Since Weyl's one-hundredth birthday is in two day's time, he must have been a mere fourteen at the turn of the century, so unless he was doing great things even at that time, I think that we can count him as being a mathematician entirely of this century! I should say, also, that as well as being such a wonderful mathematician, Weyl made several very profound contributions to theoretical physics. He was, I think, quite unusual, among pure mathematicians, in the extent and the great depth of his insights into the workings of the physical world.

Professor Yang, in his lecture, has extensively covered several of the most influential of Weyl's contributions to physics; and likewise Professor Borel will be dealing with the part of Weyl's work which has had the most impact on mathematics. I am not altogether sure where that leaves me. There are, of course, many other things in Weyl's mathematical output which were of great significance and profundity – and he had influence also in other areas such as logic, philosophy and perhaps even art. I could, I suppose, have chosen to base my own lecture on any one of these. There will, indeed, be relevance to what I have to say in Weyl's contributions to some of these areas. But rather than selecting one such topic, I shall use the fact that others will have dealt more systematically with Weyl's major contributions as an excuse for my being allowed to pay my respects while indulging in some flights of fancy of my own.

I must apologize for this, but it would be hard for me to present my appreciation of Weyl in any other way. I would not be good at giving, say, a

* A Hermann Weyl Centenary Lecture, ETH Zürich, November 7, 1985

survey of some particular topic or a historically accurate appreciation of the impact of some of Weyl's work. In fact I shall often simply take the fact that Weyl's name has become attached to a concept as the indication of Weyl's seminal influence – despite the fact that I know that with some other names that would be a risky presumption! Indeed, I have many times been struck by the numerous and varied ways that Weyl's name so often comes up in association with ideas which have strongly influenced my own research. In much of what I shall say, I shall indeed refer to ideas which I have been involved with myself. No doubt this is a most inappropriate way to express my appreciation of another, and I ask that my audience forgive me for such indulgences. But perhaps Weyl would not be displeased that the influence of his work continues to be so strongly felt through research that some others are now actively pursuing.

Symmetry and Patterns

I have mentioned that Weyl had an interest in art. The interplay between beauty and truth is something that held great fascination for Weyl. It is clear from his writings that he believed that aesthetic qualities are of profound importance in mathematics, and that truth and beauty are profoundly intertwined. He was once asked where his sympathies would lie whenever truth and beauty might come into conflict – to which he replied: beauty, every time (see Dyson 1979).

Weyl wrote a well-known book entitled "Symmetry" (Weyl 1952). This was not really a book on art or aesthetics, but it brought together many ideas from the worlds of art, mathematics, physics and biology, and unified them in terms of the concept of symmetry. There is perhaps also a hint of a subtle interplay between the roles of symmetry and *a*symmetry in art in Weyl's book. Figure 1, which is taken from the book – a photograph of a Greek statue from the 4th century B.C. – shows a work of art with a magnificent symmetry, yet it is as much the slight deviations from exact bilateral symmetry which give this work its particular artistic qualities. (Later on, I shall refer to one of Weyl's contributions to physics for which significance lay in the fact that it introduced a subtle asymmetry into the background of previously known symmetrical physics.) Figure 2 illustrates something taken from the world of biology. It illustrates star-fish and other creatures, each exhibiting a remarkable fivefold symmetry. This is to be contrasted with the sixfold symmetry exhibited by the snowflakes of Fig. 3 – taken from the world of physics. Indeed, there is a standard argument, elegantly recounted in Weyl's book, which shows that for any crystalline structure (as with the snowflakes just referred to) twofold, threefold, fourfold, or sixfold symmetries can occur, but fivefold symmetries are forbidden. There is an incompatibility between the exact translational symmetry, which a perfect crystal must possess, and a fivefold rotation axis. So it seems

 Hermann Weyl, space-time and conformal geometry. *Hermann Weyl, 1885–1985.* Ed. K. Chandrasekharan, 23–52. 1987

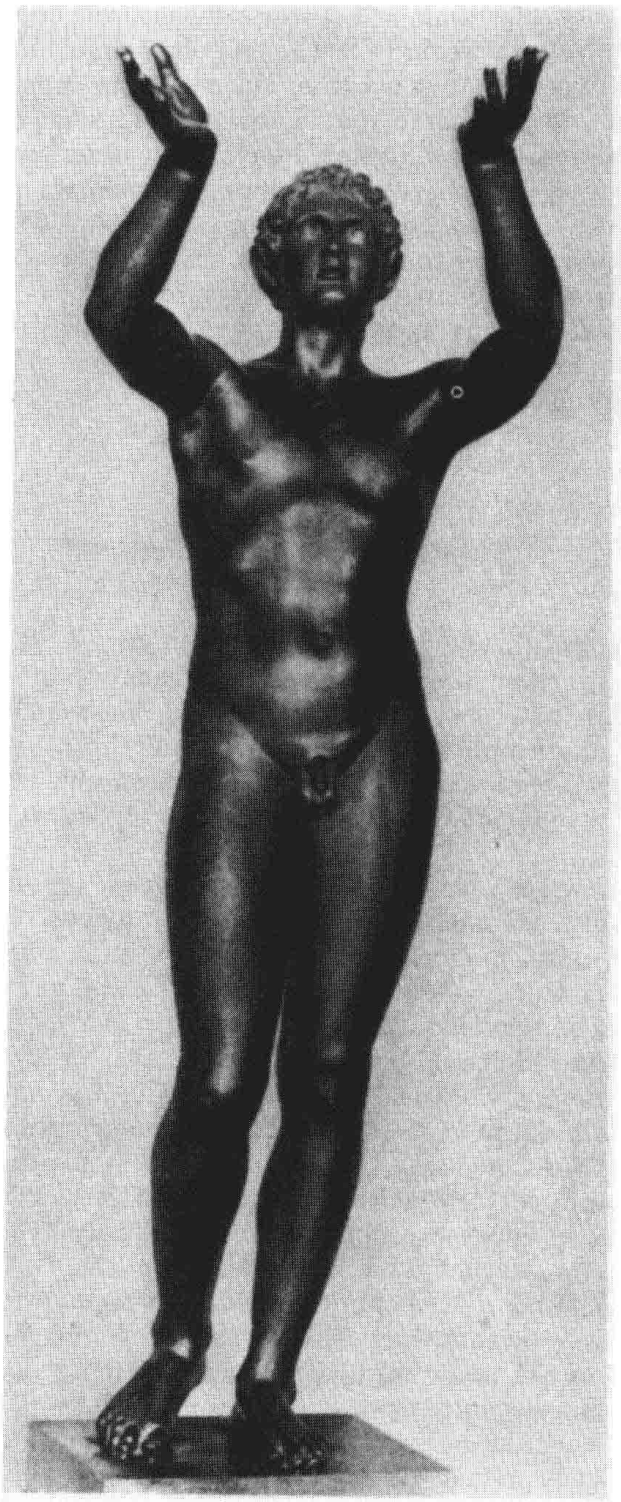


Fig. 1. Greek sculpture, fourth century B.C.

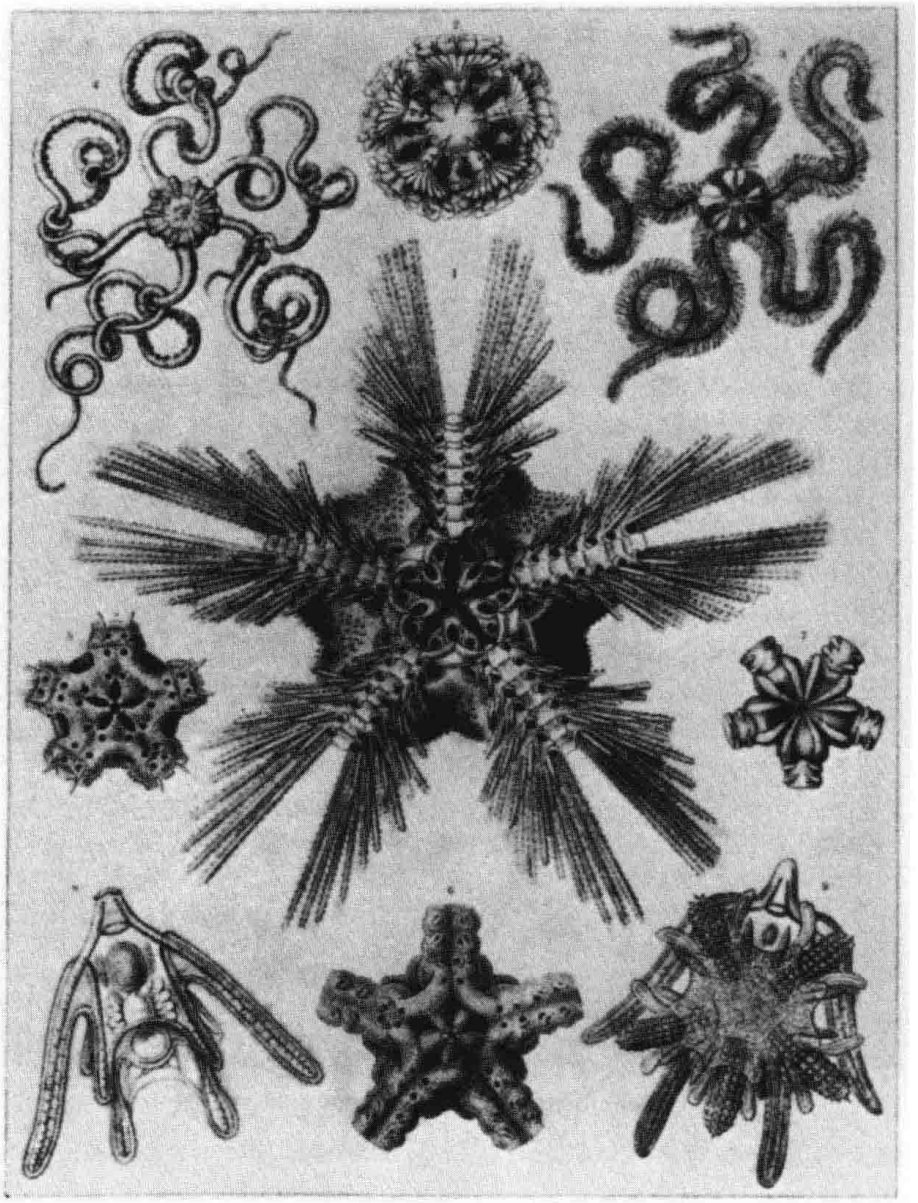


Fig. 2. Fivefold symmetry among lower animals

that while fivefold symmetry may be encountered in the world of art or of biology, we are not to find it in the crystalline structures of physics. Weyl also exhibited crystal-like symmetries taken from the world of art – many islamic designs, and, in particular, a beautiful fourteenth century window in a mosque in Cairo (Fig. 4).

Now have a look at Fig. 5. Though at first sight it seems to have the regularity of a crystalline structure, closer examination reveals that it is not

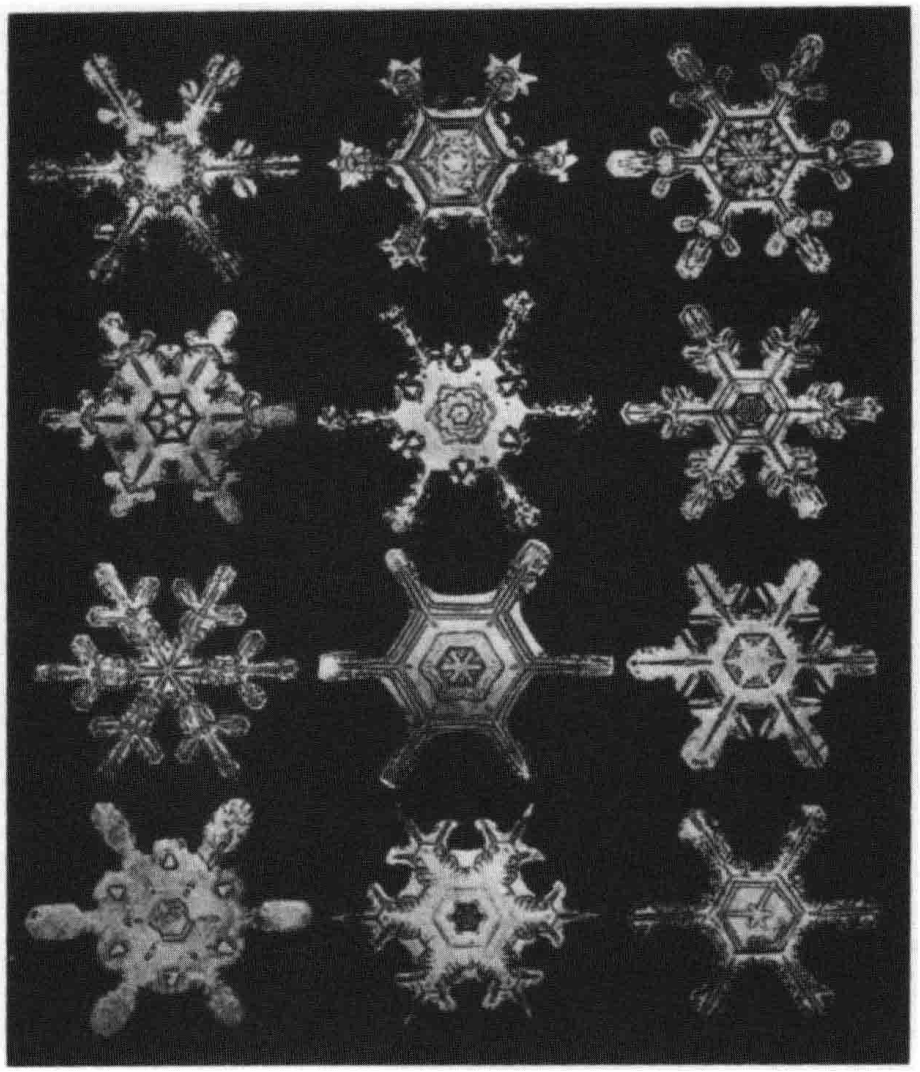


Fig. 3. Sixfold symmetry in snowflakes

exactly repetitive – and, moreover, it seems to exhibit strong elements of the forbidden fivefold symmetry! This is not an illustration taken from Weyl's book. The configuration was not known about at the time. It is something I came across somewhat over ten years ago, which, on a large scale, has many features in common with crystalline arrangements (Penrose 1974). Yet the completed pattern has infinitely many points about which it has approximate fivefold symmetry – to any preassigned degree of accuracy short of perfection. The pattern is made up of four different shapes: regular pentagons, rhombuses, pentagrams (i.e., five-pointed stars) and three-pointed "half-stars" – which I shall refer to as jester's caps. Such an arrangement of shapes covering the plane without gaps and without overlapping is referred to as a *tiling*. There are, of

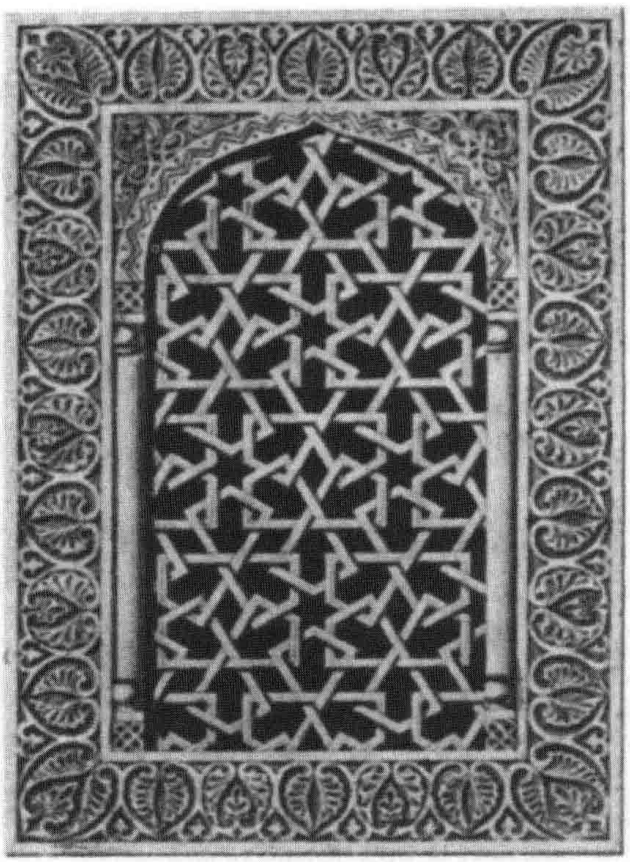


Fig. 4. Window of a Cairo mosque

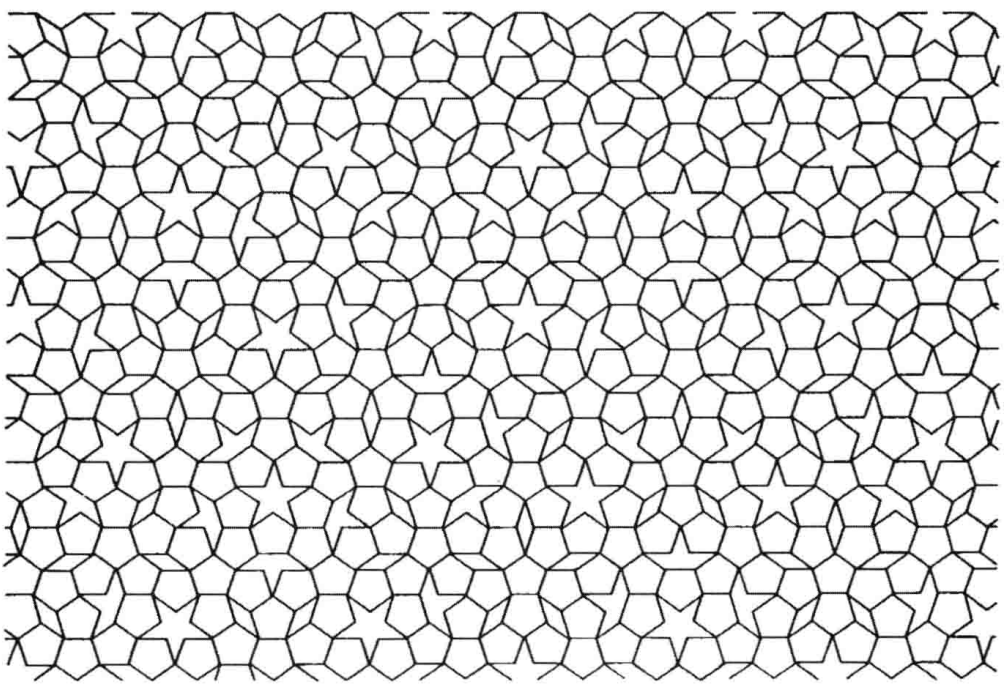


Fig. 5. Pattern with fivefold quasi-symmetry

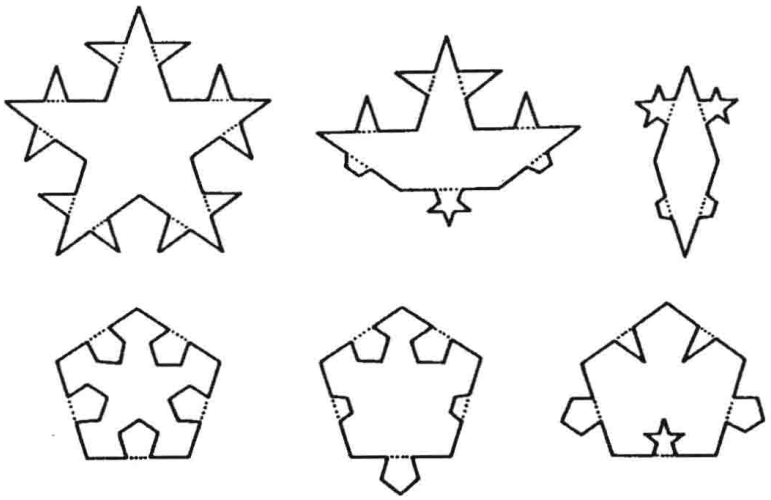


Fig. 6. Six tiles which can be assembled only according to the pattern of Fig. 5

course, many other ways of tiling the plane with these particular shapes. In particular, we could simply dispense with all the shapes save the rhombus, and tile just with that one shape in the standard periodic way. However, we can actually *force* the type of non-periodic tiling shown in Fig. 5 merely by modifying the shapes of the pieces slightly, according to the scheme shown in Fig. 6. There are now six different shapes, the pentagon occurring in three different versions according to whether it is to be adjacent to five, three, or just two other pentagons. The structure of the resulting assembly turns out to be necessarily hierarchical in nature. The nature of this hierarchy is illustrated in Fig. 7.

There are, however, many regular and uniform features that the pattern possesses which seem to go beyond its hierarchical construction. In particular, you will notice that there are many regular decagons in the pattern, each time the decagon being divided up in exactly the same way into three pentagons, two rhombusses and a jester's cap. Sometimes two decagons overlap with one another, having a single rhombus in common. Moreover, wherever there is one such decagon, it is always surrounded with a complete ring of ten pentagons. This happens even when two decagons overlap, the two rings of pentagons simply passing through one another. A further striking feature of the pattern is the way things line up in a very precise way. This is most easily seen by looking at the pattern along the page at an oblique angle. Each line segment is seen to line up with a large number of others, and in fact this feature continues indefinitely. In a certain sense the pattern is built up from strips which are almost repetitive, somewhat resembling a crystalline structure. But the strips occur in five different directions and are angled to one another at multiples of $36° (=\pi/5)$, which is crystalographically impossible! Although the pattern has translational motions that can be chosen to be as close as one likes to leaving

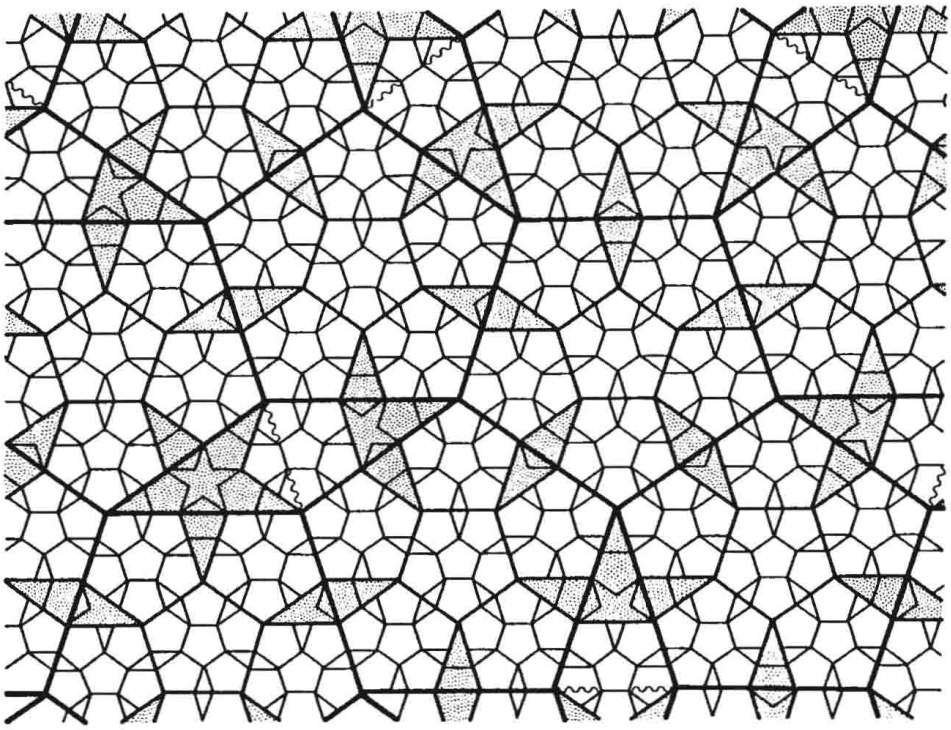


Fig. 7. The hierarchical nature of the pattern of Fig. 5

the completed pattern unchanged, none of these motions is actually an exact symmetry.

The set of six tiles depicted in Fig. 6 thus has a remarkable property: its shapes will tile the plane *only* in a non-periodic arrangement. But it was by no means the first set of tiles with this property. In fact, I should say something about the history of this problem. Suppose that we are given a finite set of polygonal shapes and we are asked the question whether or not these shapes will tile the entire Euclidean plane. Is there a *decision procedure* (i.e., a computer program which will provide a definitive answer "yes" or "no" in each case) for this problem? In around 1961, Hao Wang showed that there would indeed be a decision procedure, if it could be shown that any set of tiles, which can tile the plane in *some* way, can actually tile it periodically (Wang 1961). I think that it may have been generally felt at that time that the answer to this question would surely be "yes". (In fact Hao Wang's original result applied to a slightly different problem concerning tiles with vertices on square lattice points, but this is not an essential distinction for us here.) However, in about 1964, and following Hao Wang's leads, Robert Berger was able to show (Berger 1966) that there is, in fact, *no* decision procedure for this problem. (The argument involved showing how one may, in effect, translate a universal Turing machine into a tiling problem.) This established the existence of a set of tiles which will tile the plane only non-periodically. The argument originally led to a set of something like 20,000 tiles

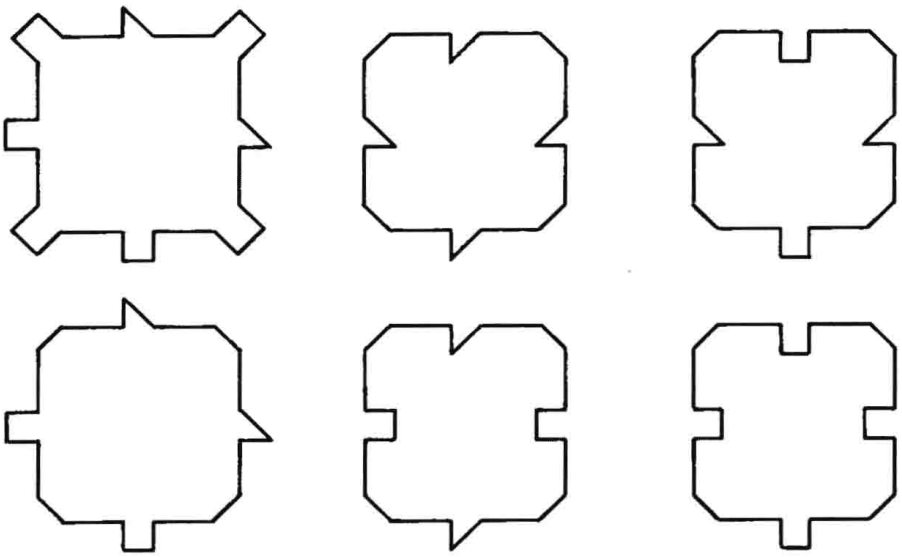


Fig. 8. Raphael Robinson's set of six non-periodic tiles

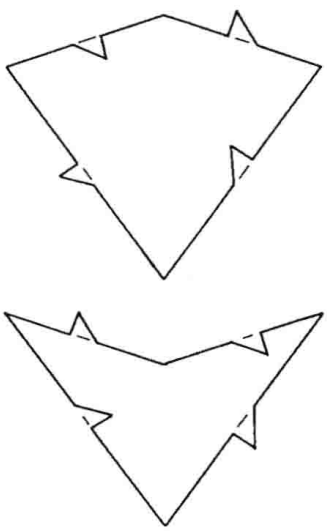


Fig. 9. Kites and darts – a non-periodic pair

with this property, though Berger was then able to reduce this number to 104. Raphael Robinson (1971) was finally able to reduce the set down to the six that I have depicted in Fig. 8. When I was told (or reminded) of this fact, it became clear to me that my own set of six could be reduced down to at most five, since the modified pentagon on the lower right can be eliminated by attaching it once to the bottom of the jester's cap and twice to the rhombus. With some more effort it proved possible to reduce the set to just two, as depicted in Fig. 9, where

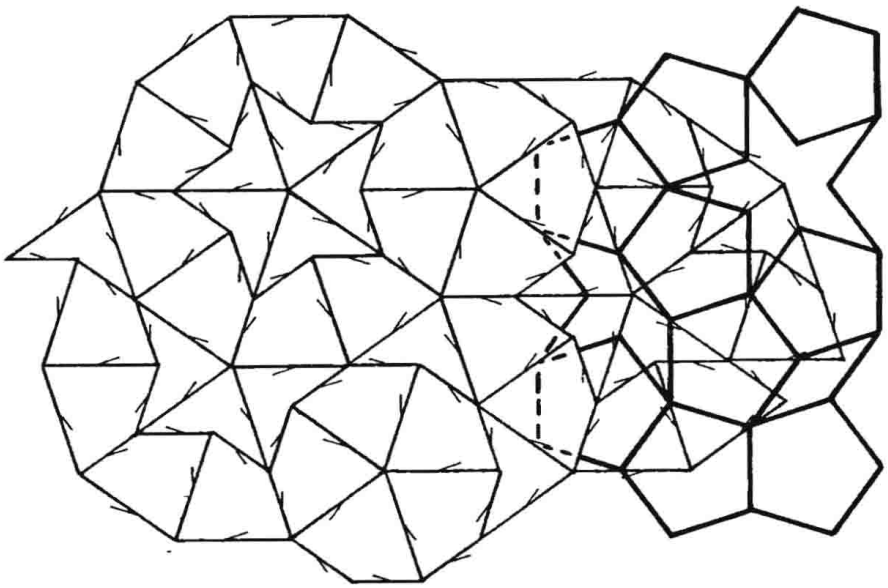


Fig. 10. Assembled kites and darts and (on the right) their relation to Fig. 5

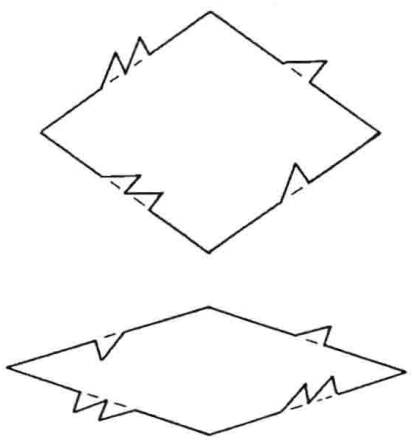


Fig. 11. A pair of non-periodic rhombuses

the tiling arrangement and the relation to my original set of six are given in Fig. 10. These two shapes are referred to as *kites* and *darts* [as described by Martin Gardner (1977) in his column in the Scientific American]. An alternative pair, consisting basically of two rhombusses, is shown in Fig. 11. An array tiled with these is depicted in Fig. 12. The pair is closely related to the original kite-dart pair and, in an appropriate sense, lies between two stages of hierarchy for the kites and darts.

There are other sets of tiles known which will tile the plane only in analogous quasi-periodic arrangements. Figure 13 exhibits a set of three tiles,

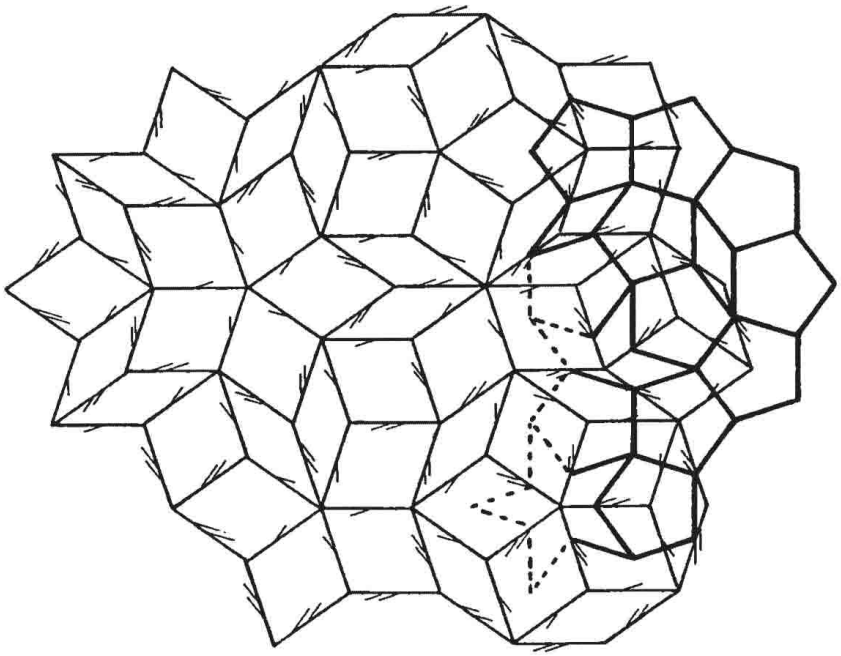


Fig. 12. Assembled rhombuses with (on the right) their relation to Fig. 5

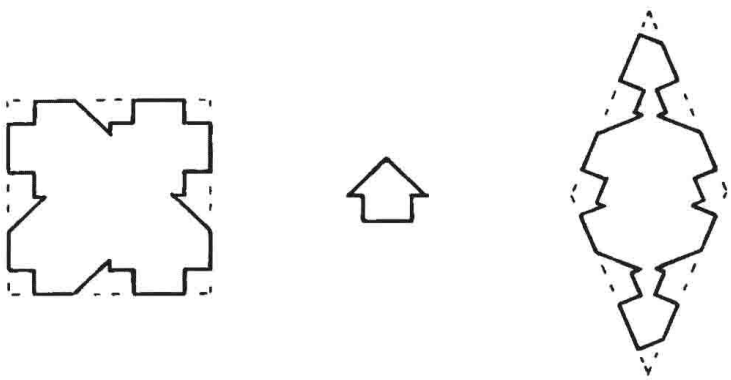


Fig. 13. A set of three nonperiodic tiles, by Robert Ammann

discovered by Robert Ammann not very long after I had found my pairs. In this case, there is always eight-fold quasi-crystalline symmetry in the completed assembly – which is also forbidden for an exact crystal (Fig. 14). There are also three-dimensional analogues of the five-fold quasi-symmetric patterns, such as an arrangement that I had found quite early on, exhibiting regular icosahedral quasi-symmetry (Fig. 15), in which interpenetrating dodecahedra meet each other along edges. The first finite set of solids which tile three-space only non-periodically appears to be a set of four discovered by Robert Ammann in 1976.

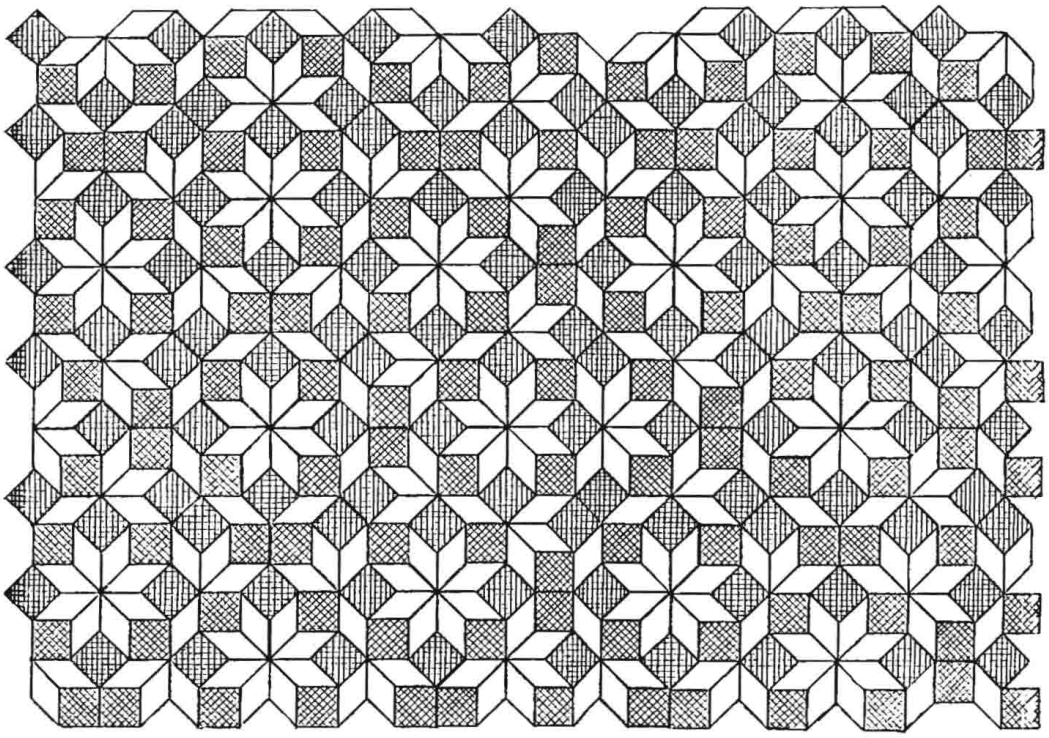


Fig. 14. An area tiled with the tiles of Fig. 13 (in essence)

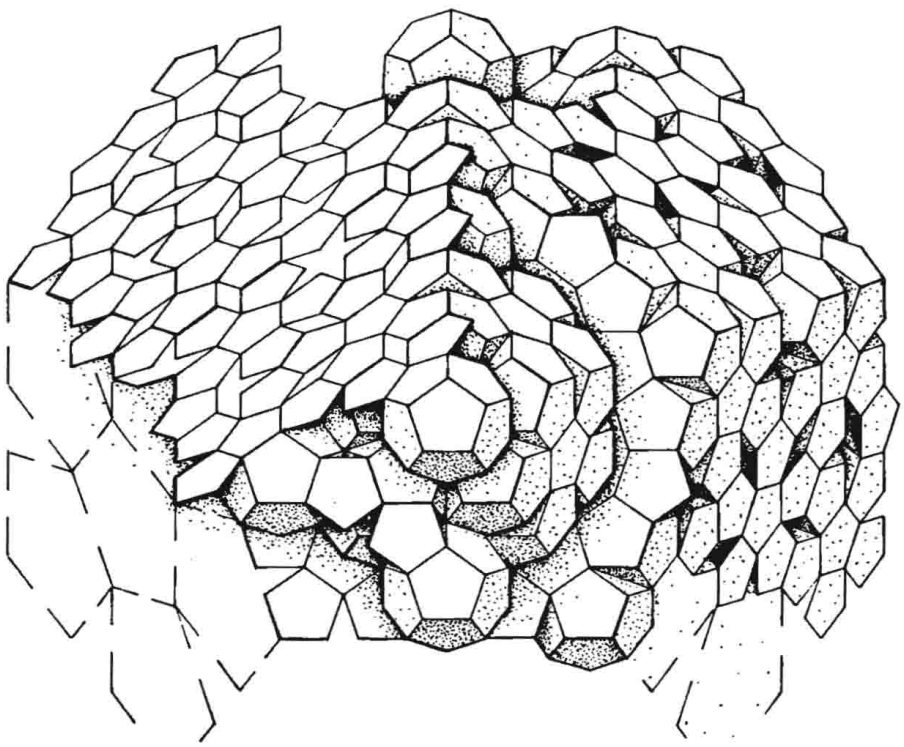


Fig. 15. A configuration in three dimensions analogous to Fig. 5

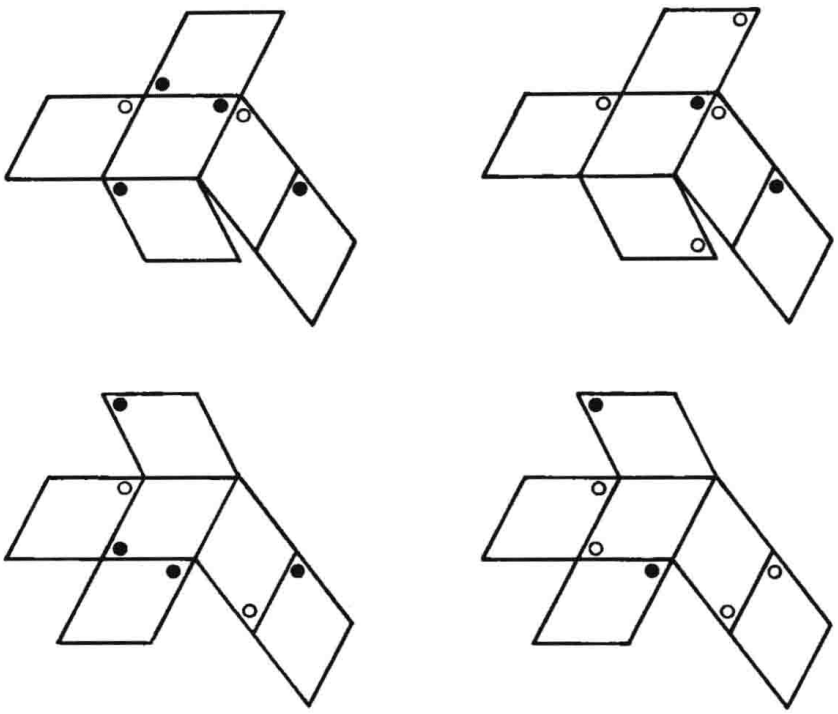


Fig. 16. Nets for assembling Ammann's four non-periodic solids; the spots must match

Nets for constructing these are shown in Fig. 16. The spots are to be matched in the assembly.

Quasi-Crystals

Suppose that the atoms of some quasi-crystalline substance were somehow arranged at the vertices of one of these patterns. How would we ascertain that fact? One method might be to obtain a diffraction pattern from it by firing electrons at it. The diffraction pattern would be, in effect, the Fourier transform of the original array of dots (vertices of the original pattern). Fourier transforms of these patterns were taken by MacKay in 1982, and later by Levine and Steinhardt in 1984. They have the striking property that they consist of sharp delta-function-like peaks (Bragg peaks) – the hallmark of an actual crystal – except that the peaks for a quasi-crystal turn out to be *dense*, and they exhibit "impossible" non-crystallographic symmetries! Remarkably, actual electron diffraction patterns from actual substances (initially aluminium-manganese alloy) were observed by Shechtman in 1983 (Shechtman et al. 1984) which have just such features. See Fig. 17 for a comparison between the experimentally observed pattern and that calculated by Levine and Steinhardt. More recently, electron micrograph pictures have been obtained of these substances

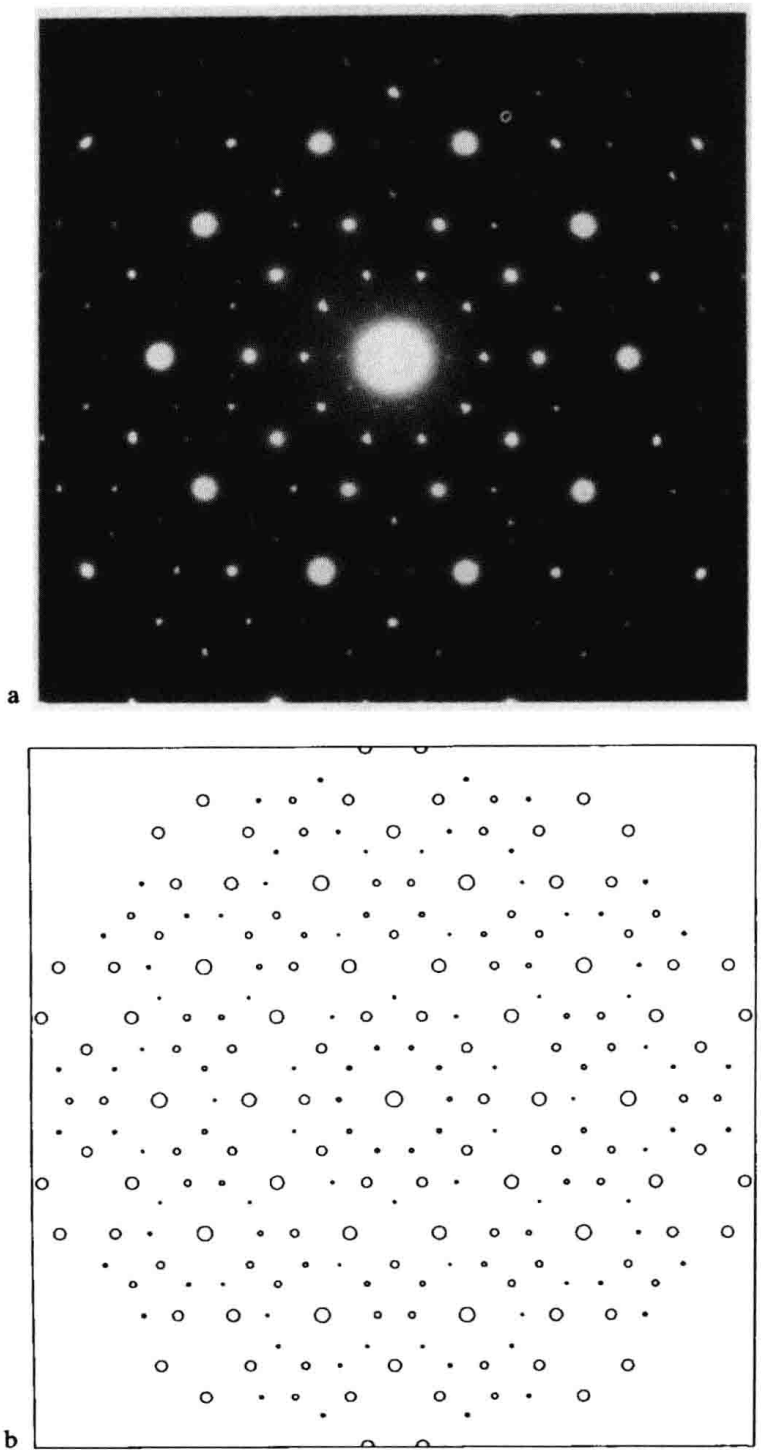


Fig. 17. An observed electron diffraction pattern (a) compared with a diffraction pattern calculated by Levine and Steinhard, 1984 (b). Both represent projections of the reciprocal structure normal to the 5-fold symmetry axis of icosahedral point symmetry. The electron diffraction pattern courtesy T. Ishimasa and H.-U. Nissen (compare Shechtman et al., 1984)

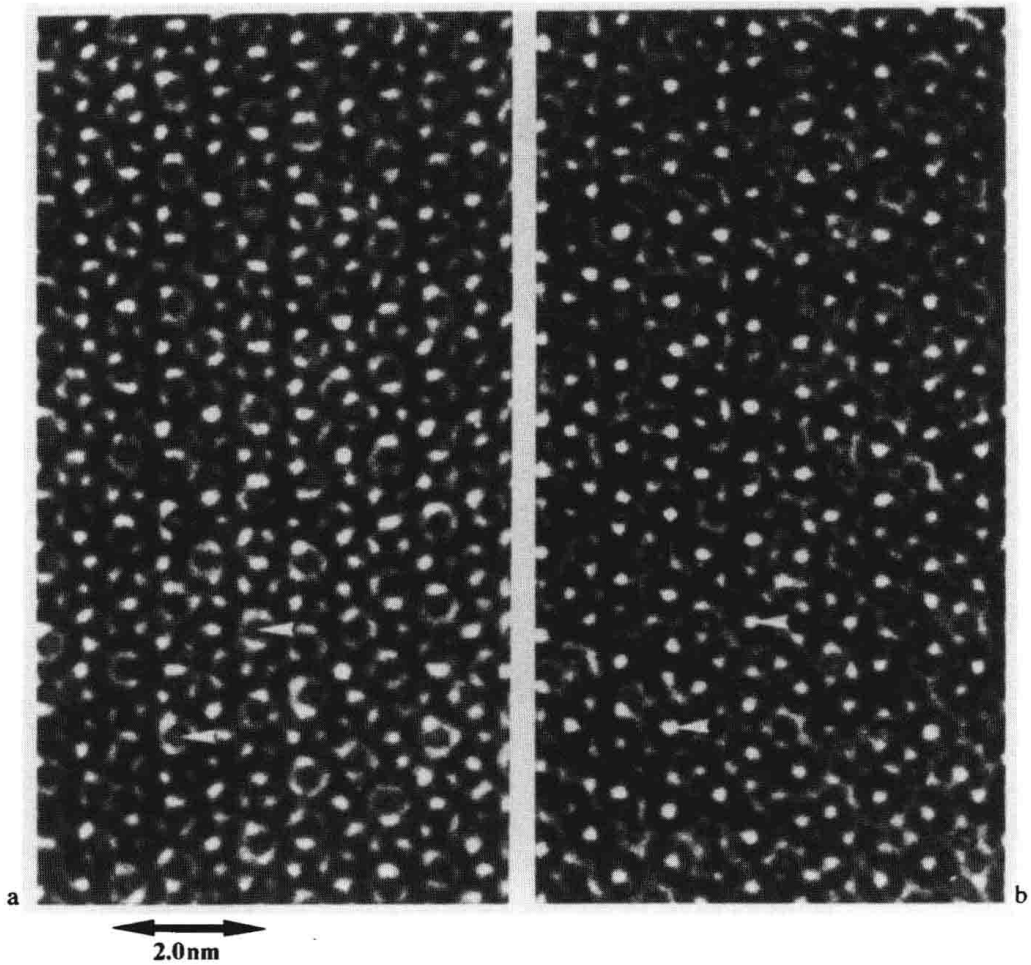


Fig. 18. High resolution electron microscopic images of the structure of Al–Mn quasicrystal projected along the 5-fold axis of icosahedral point symmetry. The objective lens of the electron microscope has a defocus of -50 ± 5 nm in (a) and -90 ± 5 nm in (b). Identical points in both images are marked by arrows. The strong bright spots in (b) correspond to the vertices of the tiling with two rhombohedrons, while these points appear in (a) as dark centres of doughnut-shaped regions. In general, for the parameters used in Fig. (a), that micrograph should give optimum correspondence with the atomic structure. Micrographs courtesy T. Ishimasa and H.-U. Nissen, ETH Zürich

(Fig. 18) which appear to give some indication of the actual atomic locations, although the exact interpretations remain somewhat obscure. Numerous pentagonal arrangements and alignments are evident, reminiscent of the patterns of Figs. 5, 10, 12. The balance of informed opinion appears to support the view that there is at least a very close relationship between such patterns and the quasi-crystalline substances[1] found in nature.

[1] There appear to be at least three different types of quasi-crystalline arrangements in nature: the icosahedral, the five-fold (or ten-fold) with periodicity in one direction, and twelve-fold with periodicity in one direction. I learned about this last type on my visit to Zurich for the purpose of giving this lecture (Nissen *et al.* 1985). One can construct hierarchial twelve-fold tilings, but I had not known of such arrangements before seeing the experimental data! I do not yet know how to force such arrangements with a finite set of tiles.

There is, as I see it, something extraordinary about this. For if one attempts to assemble such sets of tiles by hand, placing the pieces one by one where they will locally fit without much regard to the global hierarchical arrangement, then one is likely to go seriously wrong and come upon configurations which cannot be continued. That is to say, the correct assembly of these patterns depends, from time to time, upon a global examination of the assembled pattern at that stage. Whereas the rules for deciding whether or not a completed pattern has been correctly assembled are indeed local, the rules for correct piece-by-piece assembly of such a pattern are *non*-local (so the mere fact that the pieces "fit" so far does not ensure that the pattern can be continued indefinitely). So, a puzzling question is: How does Nature do it? This puzzle should be kept in mind. I shall return to the matter at the end.

Weyl Space-Times

Now I wish to turn to some completely different matters. What I have had to say so far has not really had a very detailed connection with Hermann Weyl or his work, though there is much connection with its spirit. What I shall have to say next will have a great deal more to do with him directly. I want to say something about space-time structure according to Einstein's general relativity. As far as I am aware, the very concept of a manifold – essential for any global understanding – is due to Weyl. Riemann's idea of a Riemann surface made a start towards this, but the general n-dimensional global concept appears to be due to Weyl. The idea of a manifold has very broad application within mathematics, and is certainly crucial to any global understanding of space-time structure (Weyl 1922).

I want also to refer to a specific solution (or class of solutions) of the Einstein equations referred to as the Weyl (or Weyl-Levi-Civita) solution (Weyl 1917, 1919*b*, Levi-Civita 1917). The space-times which come under this heading are the static axi-symmetric vacuums (i.e., with vanishing Ricci tensor) where some axi-symmetric source (singularity) can be present. To construct such a solution, one has just to solve the axi-symmetric Laplace equation with appropriate sources (and then apply quadratures). A particular example is obtained from the solution of the Laplace equation describing the field of two mass points at a distance along the axis. The space-time that one obtains has an interpretation as a pair of masses propped apart by a rod. One can also re-interpret the very same metric expression in a different way, namely as two masses held apart by two fishing lines extending to infinity (and also in various other ways). How can it be that the very same metric expression can have two so very different interpretations? The answer lies in how the angular variable is identified about the axis (Fig. 19). In the case of the rod, the identification is chosen so as to maintain regularity on the axis in the region beyond the masses, and this leads to a "conical" singularity along the rod; while, in the case of the fishing lines, a

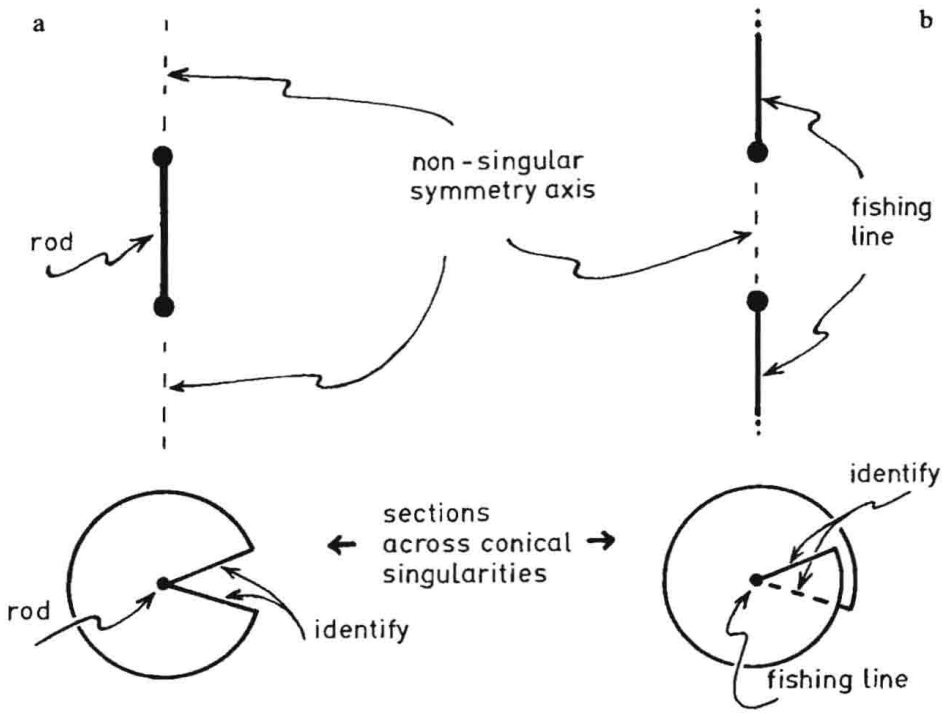


Fig. 19. The same Weyl metric represents two masses (a) propped apart by a rod (b) held apart by two fishing lines

different angular identification is chosen, which makes the axis between the masses regular but introduces a "conical" singularity along the portions of the axis extending beyond the masses (cf. Zipoy 1966). There are also various other examples in general relativity theory of the same type of phenomenon (e.g., the "C-metric" of Levi-Civita, cf. Kinnersley and Walker 1970).

One reason for my mentioning this phenomenon, is that it recently occurred to me that some structure of this kind may supply the solution to a problem that has been worrying me for a long time. The problem arose in connection with twistor theory, but its exact nature or significance will not concern us here. What the problem required was a way of "exponentiating" a relative sheaf cohomology group element to obtain some sort of deformation of a complex manifold. It has been known for a long time (Kodaira and Spencer 1958) that ordinary deformations of complex manifolds can be obtained from the exponentiation of ordinary sheaf cohomology group elements – in fact elements of $H^1(X, \Theta)$, where Θ is the sheaf of holomorphic vector fields on some complex manifold X. (Incidentally, the original idea of cohomology itself seems to have been largely another of Weyl's innovations!) What is needed here is some analogous procedure for exponentiating elements of a *relative* sheaf cohomology group – explicitly, elements of $H^2{}_Y(X, \Theta)$, where Y is a closed subset (normally a submanifold) of X. In the case of $H^1(X, \Theta)$, we may think in terms of a Čech representative. This provides us with a holomorphic vector field on

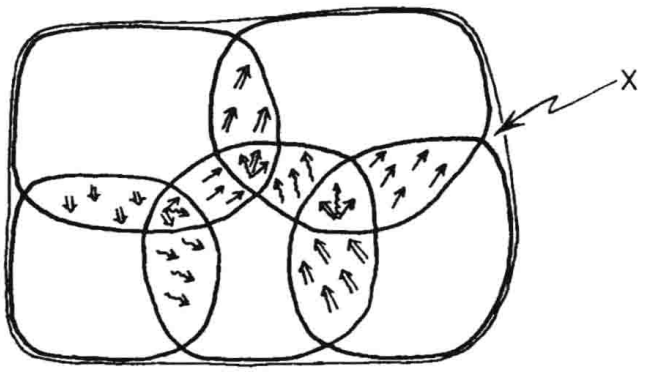


Fig. 20. Using an element of $H^1(X, \Theta)$ to deform a complex manifold

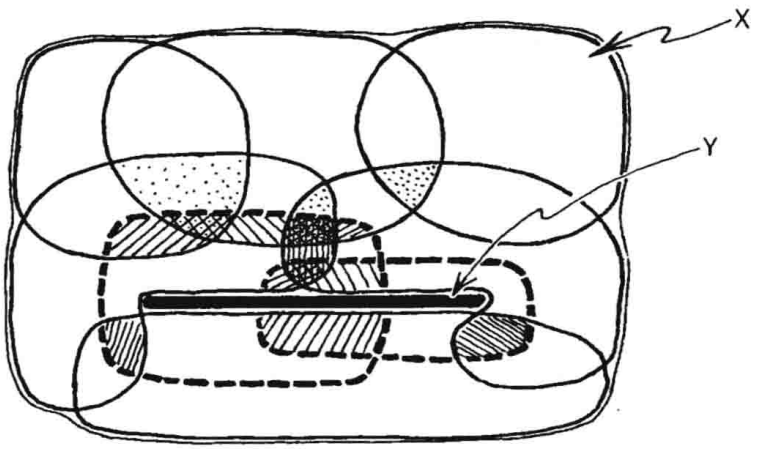


Fig. 21. A Čech representative for $H^2_Y(X, \Theta)$; vector fields defined on triple overlaps, but zero unless a set from the covering of Y is included

each intersection of each pair of open sets of some covering of X. By displacing the sets of the covering relative to one another by finite amounts along the vector fields, we obtain the required deformation (Fig. 20). In the case of $H^2{}_Y(X, \Theta)$, a Čech representative provides vector fields on *triple* overlaps of open sets of a covering of X, where at least one of the sets must in each case belong to a given covering of some neighbourhood of Y in X. A similar displacement to that given above, of the open sets along the vector fields, now gives a deformation not of X but of a neighbourhood of Y in X (Fig. 21). Thus, we have a local deformation at Y which does not extend to the whole of X, reminiscent of the local deformation of the strut in the above Weyl space-time which replaces the strut by a non-singular region, but which region does not extend to the whole of the space-time.

 Hermann Weyl, space-time and conformal geometry. *Hermann Weyl, 1885–1985.* Ed. K. Chandrasekharan, 23–52. 1987

Spinors

My next topic is spinor theory. I gained my first real understanding of spinors from lectures given in Cambridge by Paul Dirac. One thing that particularly caught my attention was a very elegant demonstration that he gave which shows that a continuous rotation through 4π can be deformed, continuously, to no rotation at all. I learnt afterwards that this demonstration was due originally to Hermann Weyl. It proceeds as follows: Consider two cones, of equal semi-angle α, one being fixed in space and the other constrained to roll on the first. When α is very small, the moving cone rolls once completely around the fixed one, returning to its starting position, and consequently executing a motion which is, in effect, a continuous rotation through 4π. We now allow the angle α to increase continuously from 0 to $\frac{\pi}{2}$, keeping the axis of the fixed cone fixed. This gives a continuous succession of closed motions. When α nears $\frac{\pi}{2}$, the motion of the rolling cone is reduced to a mere wobble. Thus we have an explicit realization of the deformation which takes a continuous 4π rotation to no motion at all.

Weyl made fundamental contributions to the theory of spinors. With Richard Brauer (1935) he co-authored a seminal paper which based the full n-dimensional discussion on a study of Clifford algebras. While the spinor concept had originally been introduced generally by Cartan, it is the Clifford algebra approach of Brauer and Weyl which is the one most followed by modern mathematicians and physicists. (There is, indeed, a considerable interest among physicists in higher-dimensional spinors.)

In a good deal of what follows I shall need to adopt the notation of two-component spinors (van der Waerden 1929, see Penrose and Rindler 1984). According to this notation, each four-dimensional vector or tensor index (denoted by a lower-case Latin letter) is replaced by a pair of two-dimensional spinor indices (denoted by unprimed and primed capital Latin letters). In a local orthonormal frame, we have a standard translation scheme for the components of the spinor translation $V^{AA'}$ of the vector V^a, namely

$$\begin{pmatrix} V^{00'} & V^{01'} \\ V^{10'} & V^{11'} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} V^0+V^3 & V^1+iV^2 \\ V^1-iV^2 & V^0-V^3 \end{pmatrix}.$$

Tensor indices are raised and lowered by using the (symmetric) tensors g^{ab}, g_{ab} of signature $(+---)$, and spinor indices by using the *skew* Levi-Civita objects ε^{AB}, $\varepsilon^{A'B'}$, ε_{AB}, $\varepsilon_{A'B'}$.

This notation enables us to write down the Weyl equation for the neutrino (Weyl 1929*a*):

$$\nabla^{AA'}\nu_A = 0,$$

where $\nabla^{AA'}$ stands for covariant derivative (a contraction taking place over the index A). Mathematicians sometimes refer to this equation as the Dirac

equation, although the usual Dirac equation for the electron couples two such fields together via a mass term. Weyl suggested that the above equation might hold for a massless particle – as it turned out, in fact, the neutrino – but this got him into trouble with Pauli, the person who first postulated the neutrino. Pauli had insisted that Weyl's equation was contrary to all the evidence of nature, for it was asymmetric under the operation of space-reflection – whereas all the evidence from atomic spectra overwhelmingly pointed to a universal symmetry in natural laws with respect to this operation. Unfortunately, Weyl did not live to witness the experimental demonstration that nature is actually not, after all, symmetrical under space reflection (Lee and Yang 1956, Wu *et al.* 1957). The Weyl equation is now an accepted equation for the neutrino!

Conformal Geometry and Conformal Tensors

Another of Weyl's contributions to physics – probably his most important such contribution – was the paper (Weyl 1918*a*; cf. also Weyl 1929*b*) in which he introduced the idea of a gauge theory. The important development of gauge theories has been thoroughly treated by Yang in his lecture, so I shall not need to discuss it here. But Weyl's work contained a host of other important ideas related to this. One of these was the idea of a *connection* in differential geometry, as something which can be defined independently of a choice of metric on the space. Christoffel and Levi-Civita had, in effect, shown how a unique connection – or, rather, a covariant derivative operator – arises from a concept of metric defined on a space; but in Weyl's original gauge theory there was only a *conformal* structure defined, so he had to allow that a more general concept of connection could arise which is not dependent upon a particular choice of metric having been made (although in his specific case he required some weaker form of compatibility with the conformal structure). The idea of a local *projective structure* and of a *projective connection* seem also to have arisen from this line of Weyl's work (see Weyl 1921).

Of prime concern for us here will be the various invariant objects that arise in *conformal geometry* – the geometry that is defined by a metric (or pseudo-metric) which is given only up to a local scale factor, so that angles are well-defined but not lengths. In the Lorentz-signature case (i.e., a space-time) the information in the conformal geometry is precisely the same as that contained in the light-cone structure of the space-time. This light-cone structure is, indeed, the physically most important part of the metric information, for it defines the causal relationships between events in the space-time. Weyl seems to have been the first to make a systematic study of conformal manifolds of higher dimension, the two-dimensional case having been treated earlier by Riemann (Riemann surfaces).

The basic invariant curvature object, in dimension $n>3$ is the *Weyl conformal tensor* $C_{abc}{}^{d}$ (see Weyl 1918*b*), where (with square brackets denoting

anti-symmetrization)

$$C_{ab}{}^{cd} = R_{ab}{}^{cd} - \frac{4}{n-2} R_{[a}{}^{[c} g_{b]}{}^{d]} + \frac{2}{(n-1)(n-2)} R g_{[a}{}^{c} g_{b]}{}^{d},$$

the Ricci tensor R_{ab} and scalar curvature R being defined from the Riemann curvature tensor R_{abcd} by

$$R_{ab} = R_{acb}{}^{c} \quad \text{and} \quad R = R_{a}{}^{a},$$

respectively. In four or more dimensions, the vanishing of the Weyl tensor is a necessary and sufficient condition for the space to be (locally) conformal to flat space. The symmetry and trace-free properties of the Weyl tensor look, at first, a little complicated:

$$C_{abcd} = C_{[ab][cd]} = C_{cdab}, \qquad C_{[abc]d} = 0, \qquad C_{ab}{}^{a}{}_{c} = 0.$$

However, these conditions are really just a statement of irreducibility under local rotations (Lorentz transformations) and describe an object which is, in the case of four-dimensional (Lorentzian) space-time, in essence remarkably simple. This simplicity is particularly brought out in the spinor formalism, where the Weyl tensor has a description

$$C_{AA'BB'CC'DD'} = \Psi_{ABCD}\varepsilon_{A'B'}\varepsilon_{C'D'} + \varepsilon_{AB}\varepsilon_{CD}\bar{\Psi}_{A'B'C'D'},$$

the 4-valent spinor Ψ_{ABCD} (with complex conjugate $\bar{\Psi}_{A'B'C'D'}$) being equivalent in information to the Weyl tensor, and *totally symmetric*:

$$\Psi_{ABCD} = \Psi_{(ABCD)}.$$

(Round brackets denote symmetrization.) This spinor always decomposes in a unique fashion $\Psi_{ABCD} = \alpha_{(A}\beta_B\gamma_C\delta_{D)}$, which leads to a simple classification scheme (known to Cartan 1922, cf. also Penrose and Rindler 1986) in terms of the coincidences between the four null directions associated with α_A, β_A, γ_A, δ_A.

Recall that Einstein's field equations tell us that the Ricci tensor of space-time is directly determined by the distribution of matter fields. The full Riemann tensor may be viewed as being a sum of two terms, one of which is the Weyl tensor and the other, a tensor with equivalent information to that in the Ricci tensor. Thus, the Weyl tensor describes precisely that part of the curvature which is *not* directly (i.e., locally) described by the matter density: the Weyl curvature describes the purely *gravitational degrees of freedom* in space-time. This turns out to be a very fruitful point of view. Rather than thinking of the metric tensor as being the object which describes the gravitational field, we use, instead, the Weyl tensor, this being an object which gives us much more specific purely gravitational information about the space-time. It is remarkable that

this information can be encoded in so mathematically elegant object as Ψ_{ABCD}, sometimes referred to as the *Weyl spinor*.

In empty space (vanishing Ricci tensor), the Bianchi identity equations become

$$\nabla^{AA'}\Psi_{ABCD}=0$$

which has a striking similarity with the Weyl neutrino equation above, and also with the spinor form of the free-space Maxwell equations:

$$\nabla^{AA'}\varphi_{AB}=0$$

(Laporte and Uhlenbeck 1931). Here, the spinor object φ_{AB} is symmetric

$$\varphi_{AB}=\varphi_{(AB)}$$

and is defined from the Maxwell field tensor F_{ab} by

$$F_{AA'BB'}=\varphi_{AB}\varepsilon_{A'B'}+\varepsilon_{AB}\bar{\varphi}_{A'B'}.$$

These differential equations are all particular examples of the source-free massless field equations

$$\nabla^{AA'}\phi_{AB\ldots L}=0, \quad \phi_{AB\ldots L}=\phi_{(AB\ldots L)},$$

the *spin* being one-half the number of indices. These equations are all conformally invariant under local rescalings of the metric

$$g_{ab}\to\Omega^2 g_{ab}, \quad \varepsilon_{AB}\to\Omega\varepsilon_{AB}, \quad \varepsilon_{A'B'}\to\Omega\varepsilon_{A'B'},$$

where we take

$$\phi_{AB\ldots L}\to\Omega^{-1}\phi_{AB\ldots L}.$$

Note that in the case of gravity (spin 2) we already have conformal invariance for the Weyl spinor. But the relationship with the above invariance is a little bit subtle because the Weyl spinor conformal invariance occurs with a different factor:

$$\Psi_{ABCD}\to\Psi_{ABCD}.$$

The Weyl spinor is a conformal object and gravity exhibits some conformally invariant properties, but gravitational theory as a whole is not conformally invariant.

 Hermann Weyl, space-time and conformal geometry. *Hermann Weyl, 1885–1985.* Ed. K. Chandrasekharan, 23–52. 1987

There is also conformal invariance for the massless field equation for spin 0, provided that we take the appropriate form of the wave equation, namely

$$\left(\nabla_a\nabla^a + \frac{n-2}{4(n-1)}R\right)\varphi = 0$$

with

$$\varphi \to \Omega^{(2-n)/2}\varphi\,,$$

where I have now reverted to a general n-dimensional description. It is my understanding that the above modified wave (or Laplace) operator and its conformal invariance are due to Weyl. Recently Eastwood and Singer (1985) have found a modification to the *square* of the wave operator, in four dimensions, which is conformally invariant, but with a different scaling from the above, namely

$$\nabla_b(\nabla^b\nabla^a - 2R^{ab} + \tfrac{2}{3}Rg^{ab})\nabla_a$$

acting on φ with

$$\varphi \to \varphi\,.$$

A new general theory of conformally invariant differential operators has been put forward by Eastwood and Rice (1986) and developed further by Baston (1986). I feel sure that this theory is something that would have been close to Weyl's heart since the methods used depend, to a considerable degree, on the representation theory of semi-simple Lie groups.

Conformal Structure of Space-times

Conformal geometry is important for general relativity also for other reasons. I have mentioned that the causal relations between events (that is, between space-time points) are determined merely by the conformal metric of the space-time and do not require the metric scaling. Moreover, the concept of a null geodesic (or light ray) is also conformally invariant. (I imagine that this fact was known to Weyl, but I do not know who first obtained it.) These properties are important for the global study of space-times. The proofs of singularitiy theorems depend, to a very large extent, on a detailed study of the causal structure of space-times and on the geometric properties of null geodesics (cf. Penrose 1972, Hawking and Ellis 1973). For a good proportion of the discussion, space-times are viewed as conformal geometries, and only now and then does the actual scaling of the metric play a role. Thinking of a space-time as

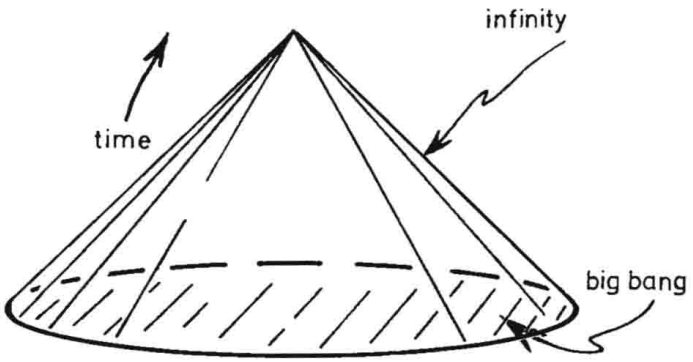


Fig. 22. Conformal representation of the open Friedmann model

a conformal geometry is also useful in other ways. In particular, it provides a method of studying the structure of *infinity* by, in effect, transforming infinity to a finite region and, as one can frequently do, treating "points at infinity" as perfectly ordinary boundary points to a smooth conformal manifold. This is very useful for the study of gravitational radiation, for example. It is often much more convenient to treat such matters by examining a little local conformal geometry than to have to consider complicated asymptotic limits (cf. Penrose and Rindler 1986).

This same kind of idea can also be used in the reverse sense. Rather than "squashing infinity inwards" to a finite hypersurface, one can sometimes take a space-time singularity and "stretch it outwards" by a similar infinite factor and transform the singularity into a smooth hypersurface with a well-defined, physically determined conformal geometry. (See Fig. 22 for an example having both types of conformal boundary at once.) In fact, this procedure only works well for certain very "regular" types of singularity. However, a singularity of this type indeed occurs at the "big bang" of a standard Friedmann-Robertson-Walker cosmological model (that is, of a spatially homogeneous and isotropic space-time). Moreover, it is a remarkable observational fact that: in our *actual* universe, there is a very close accord with such models, and this agreement gets better and better the farther we explore back in time (i.e., the closer and closer we get to the big bang). Thus, this procedure is very appropriate for studying the early structure of the actual universe. The procedure yields particular insights into the nature of cosmological horizons.

The Origin of the Second Law of Thermodynamics

In the reverse direction in time, we have singularities inside black holes – and also possibly at a "big crunch", the final all-embracing space-time singularity possessed by some cosmological models. However, such *final*

singularities would not be expected to have the regularity of conformal structure that the *initial* big bang apparently had. In a final singularity the Weyl curvature would be expected to diverge to infinite values; whereas in the initial singularity the Weyl curvature seems to have been zero (or, at least, very small). This distinction is related to the *entropy* content of the singularity, and to the *Second Law of Thermodynamics*. The Second Law states that the entropy of the universe increases with time. Thus, if the universe begins and ends with a singularity, the entropy in the initial singularity must be much lower than that in the final singularity. The distinction lies in the type of geometry which can occur near the singularity. As far as the *matter* content of the universe is concerned, one can assume effective maximization of the entropy both at the initial and final singularities. At least, that is the way that the calculations are performed at the big bang: thermal equilibrium for the matter is assumed (superimposed on an expanding background), and some excellent agreement with observational fact is obtained on the basis of that assumption. The lowness of the initial entropy was not a feature of the matter content of the universe, but was a feature of the very special *geometry* that occurred at the big bang. In effect, the *gravitational* degrees of freedom – i.e., the Weyl curvature components – were (for some unknown reason) set at zero (or small) in the initial state. The big bang seems to have been a remarkably "regular" type of singularity, of closely Friedmann-Robertson-Walker form, with zero (or extraordinarity small) Weyl curvature.

Owing to the smallness of the gravitational coupling constant, it takes a long while before the gravitational degrees of freedom – as measured by the Weyl tensor – can themselves become "thermalized". Only gradually, as the initially uniform matter in the universe begins to clump together owing to its mutual gravitational attraction, does Weyl curvature begin to appear, largely in the spaces between the clumped regions of matter. Galaxies are formed with their stars – those condensed lumps of matter which become "hot spots" of thermal imbalance. One of these hot spots (our sun) provides the thermal imbalance (or entropy sink) that is taken advantage of by the plant life of this earth. We, in turn, take advantage of the low entropy in the plants in order to live.

It is important to realize that it is *gravitational clumping* that enables the sun to shine. Nuclear reactions actually slow down the sun's inward collapse and maintain the sun's temperature at a value conducive to life. Entropy increases in our solar system, just as it does everywhere else. Life merely takes advantage of the comparative lowness of the entropy of the sun – and the lowness of that entropy is due to the fact that the sun has been able to call upon the low gravitational entropy of the uniform matter distribution from which the sun condensed. In a gravitational context, the uniform distributions (small Weyl curvature) tend to have low entropy, and the more clumpy that the distribution becomes (i.e., with larger and larger Weyl curvature) the higher will be the gravitational contribution to the entropy. (This is owing to the universally attractive nature of the gravitational field.) In a collapse to a black hole,

absolutely enormous values of entropy can be achieved (according to the Bekenstein-Hawking formula, cf. Wald 1984). Even larger values can be achieved in a big crunch – the "ultimate black hole". The final singularity has a complicated Weyl curvature diverging to infinity.

Since it is the high entropy states which are to be regarded as "probable" states, the mystery that the Second Law presents us with is: why did the universe start off in such a "highly improbable" initial state? We should also take note of the *particular* way in which it seems to have been "improbable" namely that the Weyl curvature started at zero (or, at least, very small). Setting aside, for the moment, the question as to what the *reason* might be, one can at least formulate a conjecture (which I have referred to as the Weyl curvature hypothesis, cf. Penrose 1986) to the effect: in any initial space-time singularity the Weyl curvature tends to zero in a parallelly propagated frame as the singularity is approached from a future direction. From this hypothesis one can, in effect, derive both the Second Law and the striking uniformity of the universe in the large[2].

Objective State-Vector Reduction?

I should like to end on a somewhat more speculative note. It is not unreasonable to assume that the physics which governs the structure of space-time singularities is some hitherto undiscovered theory of *quantum gravity*. Since the Weyl curvature hypothesis refers to initial singularities and not to final ones, such a quantum gravity theory would have to be *time-asymmetric*. Now this is not a conventional view amongst theorists. But in my opinion, it is a necessary ingredient in any ultimately successful theory of quantum gravity. In my view, also, the successful theory would need to involve a change in the very rules of quantum mechanics. There is, indeed, something very unsatisfactory about the strange hybrid way in which the standard rules of quantum mechanics must be applied, with two mutually incompatible modes of evolution of a physical system: deterministic unitary (say Schrödinger) evolution on the one hand – so long as no-one is "looking" – and probabilistic non-unitary state-vector reduction on the other, whenever an "observation" is being made. In my opinion, some more comprehensive and uniform evolution procedure is required which would incorporate both of these different types of evolution as limiting cases.

In the absence of a theory of such an evolution procedure, one requires, at least, some objective criterion for determining those circumstances under which unitary evolution is going to be an excellent approximation to the truth, and

[2] There is a currently popular viewpoint that this uniformity can be deduced as a consequence of the "inflationary model of the universe". However, this is a fallacy, and some equally "inprobable" starting assumptions must be postulated before such a time-asymmetric deduction can be made.

those circumstances under which state-vector reduction is closely approximated. In my view, the distinction between these two cases lies in whether *gravity* – in the sense of sufficient change in the Weyl curvature – is effectively absent or present in the process under consideration.

One clue to this can be gleaned if we consider the implications of the Weyl curvature hypothesis. Suppose we imagine an isolated container with reflecting walls, which has in it a large amount of matter. At certain times in its history, there will be a black hole in the container, while at other times not. We consider the phase space of the contents, and divide it into two regions B and A according to whether there is, or is not, a black hole in the container. The Weyl curvature hypothesis is assumed, and this precludes "white holes" – the time-reverses of black holes – because their singularity structures are unsuitable; and it also precludes space-time singularities of all other types than those arising from black holes. Black holes can form by a process of classical collapse, but they can only disappear via the quantum process of Hawking evaporation. This has the implication that there are fewer flow lines leaving the region B than there are entering it (Fig. 23). This entails a violation of Liouville's theorem (or its quantum analogue) in region B since there must be destruction of phase-space volume (convergence of flow lines) in its interior. This is not really so surprising, because information is lost at the black hole's singularity. It is a loss rather than a gain because the singularity is of "final" rather than "initial" type – the latter being generally excluded because of the Weyl curvature hypothesis.

There is, however, a problem with regard to region A. Here we have *creation* of phase-space volume (production of new flow lines) to compensate for the loss in B; but there are no black holes or singularities in the container when the phase-space point is in A! What, then can be responsible for this increase in flow

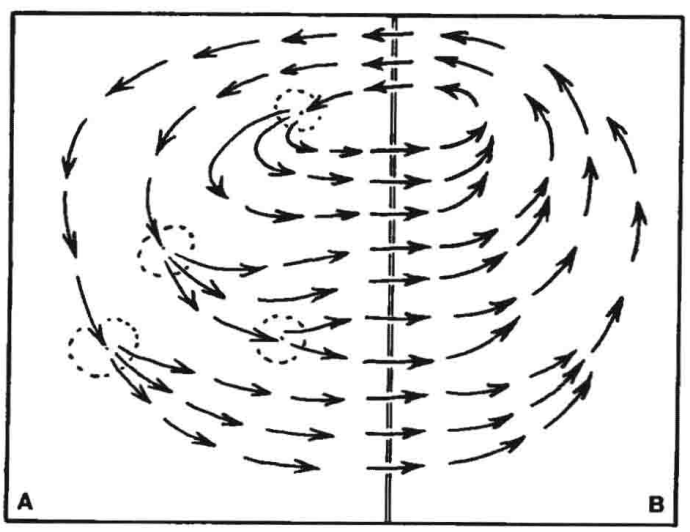


Fig. 23. Phase space for isolated container; phase-space volume lost in the black-hole region B is regained in A – perhaps because of objective state-vector reduction

lines at places in region A, where there appears to be no exotic physics taking place? My suggestion is that this is the phenomenon of quantum state-vector reduction. For state-vector reduction is probabilistic, not deterministic: different outcomes can result from the same input. Accordingly, several different flow lines in phase-space can branch out from a single flow-line input. If this contention is correct, we see that state-vector reduction is "the other side of the coin" from the Weyl curvature hypothesis, and that, indeed, state-vector reduction must be a *gravitational phenomenon.*

Let us try to see whether this is at all plausible from the point of view of orders of magnitude. The idea will be that the linear superposition of states will cease to be maintained by nature as soon as the states become significantly differently coupled in to the gravitational field. Imagine two different distributions of mass-energy initially occurring as part of a quantum linear superposition. When the gravitational fields of the two distributions become significantly different, I would contend, then linear superposition fails, state-vector reduction takes place, and one or the other of the two states is singled out. What do I mean by "significantly different"? I mean that the difference between the two Weyl tensors (which can be assumed, at this approximation, to be linear fields on a flat space-time background) is a spin 2 field whose *graviton number count is at least one graviton.*

Let us consider a particular example. Imagine a symmetrical atom which decays, emitting an electron. As the electron leaves the atom, its wave function is a spherically symmetric wave – a linear superposition of all the possible straight-line paths outwards from the atom. Let us suppose that some way out from the atom, and surrounding it, there is a cloud-chamber particle detector. When the detector is entered, each possible electron path causes a streak of ionization. These different streaks again occur all at once, in linear superposition. The various centres of ionization begin to collect droplets around them. But when each superposed track of droplets reaches a certain size, its gravitational field will become significant – and at that point, I am suggesting, state-vector reduction objectively takes place, and the reality of the world becomes just one of the tracks, rather than the linear superposition that it had been just a moment earlier. (Somewhat related ideas had been put forward many years ago by Károlyházy 1966, 1974.)

At what droplet size should this happen, according to the ideas that I am proposing? I have not yet been able to make an estimate for a whole track of droplets, but let us simplify the situation and consider just a single droplet. As an idealization, I assume that a spherical region of radius R condenses down to a symmetrical droplet of radius r leaving a vacuum between the spheres, the mass of the droplet being m. There is Weyl curvature only in the vacuum region, and one can estimate the graviton number arising, the formula being

$$768\pi^3 G\hbar^{-1}c^{-3}m^2\log(R/r),$$

where G is the gravitational constant, c the speed of light and $\hbar$ is Planck's constant divided by 2π. (I am indebted to A. Ashtekar for supplying the

numerical coefficient and confirming the logarithmic expression.) This is of order unity, for a log term of order unity, when

$$m \approx 10^{-7} \text{ grams},$$

which is not totally unreasonable.

Let me now return to the "puzzle" that I referred to earlier concerning the growth of quasi-crystals. If, indeed, they *are* accurate assemblys similar to the pentagonally quasi-symmetric tilings described in Figs. 5, 10, and 12 (and this is still not totally clear), then there appears to be something essentially *non-local* about their formation. I should like to suggest that this formation is something essentially quantum-mechanical. Rather than the atoms being laid down one-by-one, we might envisage a quantum superposition of vast numbers of different arrangements. One of these arrangements – say a quasi-crystalline one – may, as a *whole*, find favour for energetic reasons. So, when the size and energy distribution becomes right for a graviton count of order unity, one of these energetically favoured arrangements would, according to the ideas suggested above, be reduced out.

There is, of course, a good deal of speculativeness in these suggestions. I hope, nevertheless, that they may provide some possible useful hints as to directions for further developments in physical theory. Perhaps, moreover, they give some idea of how varied and significant are the implications of just one of Hermann Weyl's concepts, namely the Weyl conformal tensor. It is merely one concept among so many that Weyl has introduced, but it is one that has been particularly valuable for my own attempts at understanding.

References

Baston, R.J. (1986) The algebraic construction of invariant differential operators, D. Phil. Thesis, University of Oxford

Brauer, R., Weyl, H. (1935) Spinors in n dimensions, Am. J. Math. **57**, 425–449

Berger, R. (1966) The undecidability of the domino problem, Mem. Amer. Math. Soc. no. **66**

Cartan, E. (1922) Sur les équations de la gravitation d'Einstein, J. Math. Pures et Appl. **1**, 141–203 (cf. p. 194)

Dyson, F.J. cited by S. Chandrasekhar, Physics Today **32**, (1979), p. 27

Eastwood, M.G., Rice, J. Conformally invariant differential operators (to appear)

Eastwood, M.G., Singer, M.A. (1985) A conformally invariant Maxwell gauge, Phys. Lett. **107A**, 73–74

Gardner, M. (1977) Extraordinary nonperiodic tiling that enriches the theory of tiles, Sci. Amer. **236**, no. 1, 110–121

Hawking, S.W., Ellis, G.F.R. (1973) The Large-Scale Structure of Space-Time (Cambridge University Press, Cambridge)

Károlyházy, F. (1966) Nuovo Cim. **A 42**, 390

Károlyházy, F. (1974) Gravitation and quantum mechanics of macroscopic bodies, Magyar Fizikai Polyoirat **12**, 24

Kinnersley, W., Walker, M. (1970) Uniformly accelerating charged mass in general relativity, Phys. Rev. **D 2**, 1359

Kodaira, K., Spencer, D.C. (1958) On deformations of complex analytic structures I, II, Ann. Math. **67**, 328–401, 403–466
Laporte, O., Uhlenbeck, G.E. (1931) Application of Spinor Analysis to the Maxwell and Dirac Equations, Phys. Rev. **37**, 1380–1552
Lee, T.D., Yang, C.N. (1956) Question of parity conservation in weak interactions, Phys. Rev. **104**, 254–258
Levi-Civita, T. (1917) ds^2 Einsteiniani in campi Newtoniani, Rend. Acc. Lincei **26**, 307
Levine, D., Steinhardt, P. (1984) Quasicrystals: a new class of ordered structures, Phys. Rev. Lett. **53**, 2477
MacKay, A.L. (1982) Crystallography and the Penrose pattern, Physica, **114A**, 609
Nissen, H.-U. (1985) Material + Struktur-Analyse, Dec. 85, p. 21; T. Ishimasa, H.-U. Nissen, and Y. Fukano, Phys. Rev. Lett. **55** (July 1985), 511–513
Penrose, R. (1972) Techniques of Differential Topology in Relativity, CBMS Regional Conf. Ser. in Appl. Math., No. 7 (S.I.A.M., Philadelphia)
Penrose, R. (1974) The rôle of aesthetics in pure and applied mathematical research, Bull. Inst. Math. Applications **10**, no. 7/8, 266–271
Penrose, R, (1979) Pentaplexity, Mathematical Intelligencer **2**, 32–37
Penrose, R. (1986) Gravity and State-Vector Reduction, in Quantum Concepts in Space and Time, eds. R. Penrose and C.J. Isham (Oxford University Press, Oxford) 126–146
Penrose, R., Rindler, W. (1984) Spinors and Space-Time, Vol. 1: Two-Spinor Calculus and Relativistic Fields (Cambridge University Press, Cambridge)
Penrose, R., Rindler, W. (1986) Spinors and Space-Time, Vol. 2: Spinor and Twistor Methods in Space-Time Geometry (Cambridge University Press, Cambridge)
Robinson, R.M. (1971) Undecidability and nonperiodicity for tilings of the plane, Invent. Math. **12**, 177–209
Shechtman, D., Blech, I., Gratias, D., Cahn, J.W. (1984) Metallic phase with long-range orientational order and no translational symmetry, Phys. Rev. Lett. **53**, 1951
van der Waerden, B.L. (1929) Spinoranalyse Nachr. Akad. Wiss. Götting. Math.-Physik Kl., 100–109
Wald, R.M. (1984) General Relativity (Chicago University Press, Chicago)
Wang, H. (1961) Proving theorems by pattern recognition – II, Bell System Tech. J. **40**, 1–41
Weyl, H. (1917) Zur Gravitationstheorie, Annalen der Phys. **54**, 117; (1919) **59**, 185
Weyl, H. (1918*a*) Gravitation und Electrizität, Sitz. Ber. Preuss. Ak. Wiss. 465–480
Weyl, H. (1918*b*) Reine Infinitesimalgeometrie, Math. Zeit. **2**, 384–411
Weyl, H. (1919*a*) Eine neue Erweiterung der Relativitätstheorie, Annalen der Phys. **59**, 101–133
Weyl, H. (1919*b*) Bemerkung über die axialsymmetrischen Lösungen der Einsteinschen Gravitationsgleichungen, Annalen der Phys. **59**, 185–188
Weyl, H. (1921) Zur Infinitesimalgeometrie: Einordnung der projektiven und konformen Auffassung, Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen Math. Phys. Klasse, 99–112
Weyl, H. (1922) Space-Time-Matter (Methuen, London)
Weyl, H. (1929*a*) Elektron und Gravitation I, Z. Phys. **56**, 330–352
Weyl, H. (1929*b*) Gravitation and the electron, Proc. Nat. Acad. Sci. **15**, 323–334
Weyl, H. (1952) Symmetry (Princeton University Press, Princeton)
Wu, C.S., Ambler, E., Hayward, R., Hoppes, D., Hudson, R. (1957) Experimental test of parity conservation in beta decay, Phys. Rev. **105**, 1413–1415; Further experiments on beta decay of polarized nuclei, Phys. Rev. **106**, 1361–1363
Zipoy, D.M. (1966) Topology of some spheroidal metrics, J. Math. Phys. **7**, 1137

Mathematical Institute, Oxford, U.K.
and
Rice University, Houston, Texas, U.S.A.

(Received 3 March 1986)

170
QUANTUM PHYSICS AND CONSCIOUS THOUGHT

Quantum implications: Essays in honour of David Bohm. Eds B. Hiley and F. D. Peat, 105–20. *Routledge, 1987*

This appears to be my first written work expressing my unconventional ideas about the physical basis of consciousness. However, there is no mention of Gödel's theorem in this article. See Chapters 185 and 187.

The version reproduced here is from my original typescript. In the published article, immediately prior to the References, was the following note, added in proof: 'I am now of the opinion that the criterion for "reduction" referred to in Section 2 should not be phrased as though it referred to a "gravitational entropy" but rather to a "longitudinal graviton number". See the postscript in reference 19 for further clarification'. Also, note that the correction to the terminology 'the omnium' on p. [4] was suggested to me by Peter Derow.

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm*. Eds B. Hiley and F. D. Peat, 105–20. *1987*

Quantum Physics and Conscious Thought

by Roger Penrose
Mathematical Institute, Oxford

There can be few physicists who have delved as deeply into the philosophical implications of their subject as has David Bohm. It is therefore with some trepidation that I am paying my respects to him on the occasion of his retirement with an attempt of my own to relate questions of philosophy to those of the foundations of physics. The viewpoint that I am putting forward here attempts to relate questions that have been two of Bohm's major interests over the years, namely a possible breakdown of quantum mechanics on a macroscopic scale, and the physical basis of conscious thought. It must be evident to the reader that in both these topics I am venturing into dangerously speculative territory. Moreover others more expert that I (such as Wigner (1961), as well as Bohm (1980) himself) have put forward strongly reasoned views which relate these topics. Yet I do feel that my own viewpoint is both well worth describing, and sufficiently different from these others that the reader may find something of value here. But my account must necessarily be somewhat trucated, for reasons of space and of time. It is my intention to expand these thoughts at much greater length elsewhere.

1. Problems with Quantum Theory

Let me begin by expressing my personal worries and beliefs about the quantum theory. First, let it be said without

reservation that the basic scheme, for all its philosophical difficulties, is an extraordinarily beautiful mathematical structure. The strength of the theory (and here I refer to standard non-relativistic quantum theory, not to quantum field theory) lies not just in the unbelieveable range and accuracy of its physical predictions, but also on the mathematical elegance of its formalism.

It is here that I have always had difficulties with most hidden-variable theories. To me, it is no help just to improve upon the underlying philosophy of quantum mechanics by the introduction of hidden variables, if the price to be paid is the sacrifice of this mathematical elegance. Yet it must also be emphasized that, in my view, the standard theory is indeed quite unsatisfactory philosophically. Like Einstein and his hidden-variable followers, I believe strongly that it is the purpose of physics to provide an objective description of reality.

However, I do not regard indeterminacy, in the ordinary sense of that word, as being necessarily objectionable. We have become accustomed, through classical physics, to a picture of the world whose future evolution is completely determined by data on an initial Cauchy hypersurface. Yet, I find nothing a priori appealing about that idea. We know from results of general relativity (cf. Penrose 1965, Hawking & Ellis 1973) that there are otherwise seemingly acceptable space-times for which no Cauchy hypersurface exists. More serious are two other objections. In the first place, the accuracy with which one needs to know the initial data in order to be certain about the qualitative development of a system is absurdly unreasonable

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm.* Eds B. Hiley and F. D. Peat, 105–20. *1987*

(Born 1927, Feynman 1965) - and easily swamped by quantum uncertainties (if that is not begging the question). In the second, there is always the problem of deciding what initial data one should be choosing. It seems to me to be of little value to know that there is some initial data containing all information if one has no rules for determining what that initial data is allowed to be or likely to be.

Considering things on the ultimate universal scale, one needs, it would seem, some theory to tell us how the universe actually started off, or at least which provides probability values for the different possible initial data sets. But if it is the latter, it seems to me that one is not really better off than with a non-deterministic theory. I see no reason to be happier about feeding probabilities into the universe's initial state than peppering them throughout the space-time, as is done in the standard interpretation of quantum mechanics. If, on the other hand, the initial state is to be determined uniquely by some new principle, then we have the problem of understanding the extreme complication and curious interplay between precisely operating physical laws, on the one hand, and the presence of apparent total randomness, on the other. The picture is not an entirely hopeless one, however. One is beginning to become accustomed to great mathematical complication, variety and apparent randomness that can arise with very simple and precise mathematical transformations, especially where iterative precedures are involved, (cf. Ott 1981 and references therein). So perhaps all the complication, variety and apparent randomness that we see all about us, as well as the precise physical laws, are all exact and unambiguous consequences of one single coherent

mathematical structure.

Such a view I would call strong determinism. My guess is that Einstein* was hinting at such a possibility in his famous phrase: "The question that interests me most is whether God had any choice in his creation of the world". Perhaps this view would represent the ultimate and most extreme 'optimistic' attitude to the goals of science. Yet strong determinism is quite unlike ordinary determinism. There is not now any question of the future state being determined by data on an initial Cauchy hypersurface. The entire future (and past) is simply fixed by theory once and for all!

I do not propose to take sides on this grandiose issue. We are clearly extremely far from such a 'theory' even if eventually a viewpoint of that kind were to turn out to be 'correct'. My immediate purpose in bringing such matters up is largely to point out that, strong determinism aside, it seems no worse to feed probabilities into the theory 'peppered' throughout the space-time than to feed them all into the initial conditions. I shall, however, return to the question of strong determinism later.

At this point I should make mention of the (Everett 1957, cf. DeWitt and Graham 1973, Deutsch 1985) many-worlds type of viewpoint. Here one may be allowed a form of 'strong determinism without determinism'. The totality of all possible universes may be thought of as a single structure - the plenum* - and one might

---

*See Pais (1982); p. ...

*This terminology was suggested to me by D. Deutsch

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm*. Eds B. Hiley and F. D. Peat, 105–20. *1987*

take the view that it is the ~~plenum~~ omnium that is completely fixed by mathematical rules. The probabilities (or randomness) now arise owing to the uncertainties involved in the question of where one finds oneself located within the ~~plenum~~ omnium. This 'location' involves not only one's spatio-temporal location within a particular universe branch, but also the selection of that particular branch itself.

Unsatisfied though I am by such a world-view, I do not have any really fundamental objection to pictures of this general kind. One problem I do have with them, however, is that the continual branching of the world and the threading of my own consciousness through it would seem to result in my becoming separated from the tracks of consciousness of all my friends. This is what I have referred to as the 'zombie' theory of the world (cf. Penrose 1980 : p. 595). It seems to me that one needs a respectable theory of consciousness before the many-worlds view can hang together as a physical theory and as a viable interpretation of quantum mechanics. This strikes me as a tall order at the present time. I shall have more to say about the many-worlds viewpoint later.

Let us now consider what the standard formalism of quantum mechanics has to say about the evolution of the world. Taking this formalism at its face value, we have a state vector $|\psi\rangle$ which evolves for a while according to the completely deterministic Schrödinger equation*. Then, at odd times, when an 'observation'

---

*If preferred, one could of course use the Heisenberg picture instead, in which case the state itself is considered not to evolve in time. The distinction is not important for my purposes. The two pictures are completely equivalent, but I feel that, at least for a non-relativistic discussion, the Schrödinger picture is less confusing.

is deemed to have been made, the Schrödinger-evolved state vector is discarded and replaced by another, which is selected in a random way, with specific probability weightings, from among the eigenvectors of the operator corresponding to the observation. As has been argued on innumerable occasions, this is a wholly unsatisfactory procedure for a fundamental description of the 'real world'. There are, of course, very many different attitudes to the resolution of this problem, and it is not my purpose here to enter into a discussion of all of them. But some brief remarks will be in order.

In the first place, it is often argued that $|\psi\rangle$ itself should not be regarded as giving an objective description of the world (or of part of it) but as providing information merely of 'one's state of knowledge' about the world. This view I really cannot accept. Quite apart from the question as to who the 'one' might be in this statement (and the 'one' is surely not me!), it seems to me to be perfectly clear that there is (if we accept standard quantum mechanics) a completely objective meaning to $|\psi\rangle$ - at least to the ray determined by $|\psi\rangle$ in the Hilbert space (so that uncertainty $|\psi\rangle \mapsto e^{i\theta}|\psi\rangle$ is not of significance). For we can, in principle (according to the theory), set up a measurement defined by the operator

$$Q = |\psi\rangle\langle\psi|$$

and find that $|\psi\rangle$ (assumed normalized: $\langle\psi|\psi\rangle = 1$) is the unique state (up to phase) corresponding to the eigenvalue unity for Q. This $|\psi\rangle$ is distinguished from all other states by the fact

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm*. Eds B. Hiley and F. D. Peat, 105–20. *1987*

that it yields the value unity with certainty for the measurement Q. This is an entirely objective property, so we conclude that (if the theory is correct) the property of being in state $|\psi\rangle$ is, indeed, completely objective.

In practice, however, it may well turn out that the actual performing of the measurement Q is quite out of the question and it is in such circumstances that I would myself begin to doubt that $|\psi\rangle$ actually describes 'reality'*. But the taking of such a view requires a denial of one of the fundamental tenets of quantum theory: any (bounded) Hermitian operator, such as Q, represents a measurement that could in principle be made.

So it seems to me that one must take it that in quantum mechanics the state vector <u>is</u> actually intended to represent 'reality'. The trouble comes, of course, when a measurement is made for which $|\psi\rangle$ is <u>not</u> in an eigenstate.

Now $|\psi\rangle$ jumps non-deterministically into some new state - this is state-vector <u>reduction</u>. Because of this lack of determinism it is not possible to consider such behaviour as resulting from the Schrödinger evolution of some larger system which includes the apparatus configuration as part of the state. (Some attempts have been made to pin the blame for this indeterminism on a lack of knowledge about the environment, the claim being that when this random environment is taken into account, Schrödinger evolution might hold always. I have

---

*There are situations, in the context of relativity or in connection with time-symmetry considerations, where the 'reality' of $|\psi\rangle$ seems to lead to a paradox (cf. e.g. Penrose 1980). I am not concerned with such matters here, but such 'paradoxes' are part of my own doubts about the complete validity of the quantum formalism.

not yet found these attempts to be very believeable and I shall ignore this possibility in the discussion which follows.) So it seems that quantum mechanics asserts two quite distinct types of evolution, deterministic Schrödinger evolution and state-vector reduction.

If we accept that state-vector reduction is a real physical process - and the physical objectivity of the state vector itself seems to imply this (leaving the many-worlds viewpoint aside) - then we may ask at what stage does such a reduction take place? Again there are different views. Perhaps it occurs quickly and spontaneously at some level just not quite reached by experiments designed to detect violations in Schrödinger evolution. Or perhaps it is delayed until the latest allowable moment, as the results of the observation finally enter the mind of some conscious observer (Wigner 1961). As we well know (von Neumann 1955) either view, or any other viewpoint according to which the reduction takes place at some intermediate stage between these two extremes, is euqally and completely compatible with all the experiments.

This frustrating and beautiful fact is, at one, among the greatest strengths and the most disturbing weaknesses of the theory. It is such a strength because it enables the theory to operate - as it does with such extraordinary accuracy and power - without one needing to have the remotest idea of what 'actually' takes place during the reduction process. This strength is also the theory's weakness since for that very reason it offers us almost no clue as to what is physically going on during reduction. Moreover, it leads us into endless arguments about 'interpretation' where strange (and, to my mind, highly

questionable) philosophy is often invoked to lull one into thinking that no new physical theory is needed to explain the details of the actual physics of reduction.

The reader will realize that I am expressing a very personal view here, and that much more discussion and open-mindedness is required than I have allowed myself in the foregoing remarks. However, my purpose here is not to be open-minded but to make some suggestions. So I hope that the reader will continue to bear with me and, instead, take upon himself or herself this burden of open-mindedness. I wish to make some comments concerning the place to look for a new physical theory of reduction, and then to make a speculative suggestion as to how this might ultimately relate to the phenomenon of consciousness. In bearing with me, the reader must, as a first step, be prepared to envisage that there is indeed a real physical process of reduction, where the Schrödinger equation and linearity are presumably both violated, and where determinism may perhaps be violated also.

It may seem that, as I am suggesting a connection with consciousness, I propose to follow the Wigner (1961) extreme, where reduction takes place only at the level of conscious thought. In fact I am not at all happy with that viewpoint and my suggestions with regard to reduction will be quite different. I share the discomfort of many others that the 'Wigner view' seems to imply a markedly different physical behaviour in our own small corner of the universe from that which would be taking place almost everywhere else, in the absence of local

conscious observers to keep things under control. Of course, von Neumann again comes to the aid of that viewpoint to ensure that no contradiction with actual observation (by conscious beings) occurs. Or so it would seem. In any case, I personally find such a lop-sided picture of a real physical universe too unattractive to be believeable.

Alternatively, as quite a number of physicists now seem to argue, perhaps one should be beyond Wigner and say that reduction fails to take place even at the level of consiousness. This leads us, instead to Everett's many-worlds picture. It seems that the main motivation behind this idea is the understandable desire for economy and uniformity of description. Rather than having to have two seemingly incompatible modes of evolution for the universe - Schrödinger's equation and reduction - is it not more economical and unified to settle just for the former and discard the latter? A faith in the elegance of Schrödinger linearity and the strong experimental support for Schrödinger evolution are among the arguments put forward for this view. But in my opinion, both arguments are somewhat misconceived. Linearity seems elegant only when one has not seen an even more attractive essentially non-linear generalization. (Compare the analogy of Newtonian gravity and general relativity.) And the unquestionably impressive experimental support that quantum theory continues to have is not for the Schrödinger evolution of a quantum state. It is for that absurd concoction of Schrödinger evolution on the one hand and state-vector reduction on the other which defines the standard Copenhagen interpretation. That is where the support lies and

Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm.* Eds B. Hiley and F. D. Peat, 105–20. *1987*

and it is that with which we must come to terms in our attempts toward an improved theory.

2. Quantum Gravity and Reduction

What we seek, I am therefore claiming, is a new theory, presumably non-linear and possibly non-deterministic, which: (a) is extremely accurately approximated, for reasonably small or simple systems, by a Schrödinger-evolving state vector; and (b) accords closely, for systems which are large or complicated, with reduction and with the standard replacement of probability amplitudes by the probabilities of actual alternative outcomes. (See Pearle 1976, 1984, 1985 for a very serious concrete proposal of this kind.)

In various other places (Penrose 1979, 1981) I have put forward arguments to support my own viewpoint that the essential new and non-linear physical input should be general-relativistic gravity, together with certain considerations about the second law of thermodynamics and the apparently strongly time-asymmetric constraints on the structure of space-time singularities. I do not wish to repeat the details of my (rather round-about) argument here. My latest thinking on this question is being presented elsewhere (Penrose 1985) and I refer the interested reader to that source for more details. In my opinion there are indeed strong grounds for believing that for a quantum gravity theory to have any chance of real success it must incorporate among its important features a precise and objective theory of state-vector reduction. (A different line of thought also leading to the conclusion that state-vector reduction is a gravitational

phenomenon had earlier been put forward by Karolyhazy (1966; see Karolyhazy, Frenkel and Lukács 1985 for an up-to-date account of these ideas; cf. also Komar 1969).

There are various reasons for thinking that the structure of quantum mechanics may well have to be modified in the context of general relativity. One of these concerns the choice of variables with respect to which the process of 'quantization' is to be conducted. In the standard non-relativistic (or special-relativistic) quantization procedure one has the canonical variables x^a, p_a of position and momentum which are made into non-commuting operators and whose (canonical) Poisson bracket relations are replaced by (canonical) commutators. It would not be valid to apply this quantization procedure instead to some other variables $\tilde{x}^a$, $\tilde{p}_a$, obtained from x^a and p_a by an arbitrary canonical transformation, again replacing the canonical Poisson bracket relations of $\tilde{x}^a$ and $\tilde{p}_a$ by canonical commutators. It might, for example, be the case that $\tilde{x}^a$ is related to x^a by some arbitrary coordinate transformation, with canonical conjugate $\tilde{p}_a$ correspondingly chosen, and then the canonical quantization procedure will not in general apply to $\tilde{x}^a$ and $\tilde{p}_a$. What distinguishes the standard (x^a, p_a) from the more arbitrary $(\tilde{x}^a, \tilde{p}_a)$ (cf. Komar 1971) is its direct relation to the symmetry group of flat space-time, so that p_a indeed represents the standard (conserved) energy-momentum concept.

It will be noted that this selection principle for 'preferred' canonical variables will not work in a curved space-time. This is a much more immediate problem than the daunting difficulties involved in trying to quantize the gravitational field itself, since it arises already when gravity is simply taken to be an

unquantized background field. Thus, even at this simplified level, the standard quantum-mechanical procedures run into difficulties. One does not, in fact, have any clear rules for a quantizing procedure, in the general case. When one is asked the question of whether the incorporation of space-time curvature into physics in any way affects the general framework and rule of procedure of quantum mecahnics, one is forced to reply (if truthful) that one does not even know what the rules are, in a general curved-space setting!

It seems that in curved space, one must resort to some 'patchwork' formalism and, at best, much of the compelling elegance of quantum mechanics is removed. It has become fashionable to argue that, in view of the abounding difficulties that have been encountered in attempts to quantize general relativity, one should replace that theory by some other which might more willingly submit to being forced into the linear framework of standard quantum theory. Yet it is not often suggested that the quantum framework might itself be forced to yield. I would not dispute that some changes in classical general relativity must mecessarily result is a successful union with quantum physics is to be achieved, but I would argue strongly that these must be accompanied by equally profound changes in the structure of quantum mechanics itself. The elegance and profundity of general relativity is no less than that of quantum theory. The successful bringing of the two together will never be achieved, in my view, if one insists on sacrificing the elegance and profundity of either one in order to preserve intact that

of the other. What must be sought instead is a grand union of the two - some theory with a depth, beauty and character of its own (and which will be no doubt recognized by the strength of these qualities when it is found) - and which includes both general relativity and standard quantum theory as two particular limiting cases.

This is all very well, the reader is no doubt thinking, but what relevance has the extremely weak phenomenon of even classical gravity - let alone the absurdly tiny and quite undetectable 'quantum corrections' that one anticipates would result from a quantum gravity theory - to the commonplace phenomenon of reduction? Whether one contends that reduction occurs early or late in the chain of alternatives - whether reduction has already physically taken place with the track in a cloud chamber or mark on a photographic place, or whether it is delayed until it affects the state of a human brain - in either case only tiny amounts of energy are involved, by the standards of gravitation theory. Yet it is my contention that it is this link with gravity which effects the apparent change in rules that takes place with reduction. When two states of differing energy distribution are linearly superposed, the slightly differing space-time goemetries that these energy distributions produce (according to general relativity) must also be superposed. This is not clear-cut matter, since as we have seen, the geometry is not just a 'physical state' but is something essentially involved in the very determination of the 'procedure of quantization' (cf. forgoing remarks concerning x^a, p_a and $\tilde{x}^a$, $\tilde{p}_a$). It is hard to see how 'quantization procedures' are to be superposed.

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm.* Eds B. Hiley and F. D. Peat, 105–20. *1987*

I would contend, therefore, that when two geometries involved in a linear superposition become too different from one another - in some yet-to-be-determined precise sense - then linear superposition fails to hold, and some effective non-linear instability sets in, resulting in one or the other geometry winning out, the result being reduction. However the criterion for deciding when such an instability becomes operative cannot depend solely on the energies or masses involved. The masses involved in cloud chamber droplets, or relevant collections of silver iodide molecules in a photograph, or configurations in a brain, would appear to be very tiny - apparently tinier than the Planck-Wheeler mass of 10^{-5}g which seems to characterize the scale of quantum gravity - which, in turn, might perhaps be smaller than the total electron mass involved in a coherent quantum state for a ~~large~~ huge superconducting coil. Some care will be needed in deciding upon the precise measure of mass-energy distribution difference that is needed to trigger off the instabilities of reduction.

My earlier arguments (Penrose 1979, 1981) have emphasized the time-asymmetric role of entropy in reduction and the gravitational origin (via time-asymmetry in space-time singularity structure) of the second law of thermodynamics. For an overall consistency of the physics involved, it is argued that gravity ought also to be critically relevant to the quantum-mechanical reduction process. But it is gravity in its role of providing a gravitational entropy that enters. In the early stages of the universe (at the big bang) the entropy content

of the matter was high. The second law arises because, at the singularity, something constrained the entropy of gravity to be extremely low, whereas the potential for entropy content of (conformally) curved space-time was enormously high.

As the universe evolves, this unused potential of gravitational entropy is gradually taken up (clumping of gas into stars, etc.) and is directly or indirectly taken advantage of by systems requiring a low entropy resovoir (e.g. plants making use of degradable photons from the sun, which is a hot spot in the sky by virtue of its gravitational clumping). The ultimate high value of entropy for a gravitationally clumped object is achieved by the black-hole state. This entropy can be given in precise terms by the Bekenstein-Hawking formula (cf. Wald 1984) and we find that for an ordinary black hole, resulting from the collapse of a, say, ten solar mass star, the entropy per baryon is some twenty orders of magnitude larger than even the seemingly huge value of 10^8, or so (taking units with Boltzmann's constant as unity), for the thermal black-body radiation left over from the hot big bang.

Even when this maximum entropy value is not achieved, the gravitational entropy can be quite sizable, though as yet no precise determination of this entropy has been suggested. In a rough way, we can say that this entropy is intended to estimate (the logarithm of) the number of quantum states that go to make up a given classical geometry. Such an entropy measure would have to be very much a non-local expression. But this is not surprising in general relativity where even energy must be given by a non-local expression (cf. Penrose and Rindler 1985).

 Quantum physics and conscious thought. *Quantum implications: Essays in honour of David Bohm.* Eds B. Hiley and F. D. Peat, 105–20. *1987*

The idea is that for 'reduction' to take place, we must be in a situation where, such as in the localization of a photon on a photographic plate, the lowering of entropy* that this 'reduction' would seem to entail must be at least compensated by a corresponding increase in the gravitational entropy, this gravitational entropy being higher for clumped localized energy distributions than for those which are spread out more uniformly. Quantum reduction must be a non-local process, and the hope is that this non-locality can be matched with the non-locality involved in the gravitational entropy concept.

It will be clear to the reader that there is much speculation and lack of precision in this picture. (For more clarification, see Penrose 1985.) But my purpose here is not to spell out in detail how the reduction procedure might work (which I cannot do, not having a proper theory) but merely to attempt to persuade the reader of the plausibility of there being some new physical process going on which has a perfectly objective character, even though we do not understand how this process works in detail. I would certainly anticipate, however, that the process is likely to defy any meaningful local description in ordinary space-time terms. But any specualtion involved in the 'details' of this process will, I suppose, pale to insignificance by comparison with the speculative aspects of what I have to suggest in the next section - where I turn to

---

*This cannot be just the 'quantum-mechanical entropy', which remains at zero so long as one uses a pure-state description, but is some more 'common sense' (though somewhat ill-defined) entropy concept which increases as more and more degrees of freedom become involved. (I am grateful to P. Pearle for illuminating discussions concerning this.)

what will (at first) seem to be a totally different topic!

3. Concerning the physics of conscious thought

There has been a great deal of discussion in recent years - since the emergence of modern high-speed computers - of the question as to whether human thought can be adequately understood in terms of computational algorithms. There is a rather prevalent school of belief (cf. e.g. Hofstadter 1979) which maintains that an algorithm (or perhaps the end-product of an algorithm) is itself all that 'counts' in determining whether 'thoughts', or even 'conscious feelings' have been achieved during the implementation of that algorithm. According to this view, the actual physical nature of the device which carries out the algorithm is of no consequence. The point of view is a strongly 'operationalist' one and springs, to a large extent, from a well-known and penetrating discussion due to Turing (1950) according to which an experimenter has to decide, merely by means of questions and answers (transmitted impersonally with, say, keyboard and display screen) which of two 'subjects' is a human being and which is a computer - the computer having been programmed to imitate human responses as closely as possible. This has become known as the 'Turing test', and the point of view to which I have referred ('strong AI') would maintain that once a computer has satisfactorily passed this test (i.e. has become able to imitate human responses closely enough to be able to 'fool' a perceptive experimenter) then all the normal terms, such as 'thought', 'understanding', 'awareness', 'happiness',

'pain', 'compassion', 'pride', etc. could be as meaningfully and as truthfully applied to the computer as to the human subject.

I do not believe that it would be strongly disputed that computer technology (both at the hardware and software ends) is some good way from ataining this avowed goal. Nevertheless, at a much more restrictive level, there are already computer programmes which can respond to extremely simple types of stories in the English language. The computer can provide answers to elementary but slightly indirect questions about the story, these answers seeming to indicate some very primitive 'understanding' of the content of the story by the computer. However, Searle (19..) has argued impressively that answers of this sort need in no way imply that any kind of 'understanding' has actually been achieved by such a mechanism since in place of the electronics of the computer, one could use human beings blindly following instructions and they need have no inkling of the 'meaning' of the story in order to carry out these instructions successfully.

It seems to me to be clear that in this essential respect Searle is correct and we have no reason to apply such a word as 'understanding' to the operations of a mechanism, electronic or otherwise, of this kind. It appears that Searle is willing to concede that at some stage computers (of present-day type, but improved in speed, storage capacity and logical design) may be able to pass the Turing test proper (let us say, by one hundred years from now) but that we would still have no reason to attribute to them such human qualities as I have mentioned. Most particularly, we would have no reason

to attribute to them any form of conscious awareness. Thus, according to Searle, the actual physical material and detailed construction of a 'computer' (i.e. the 'hardware') may, for such qualities, be of 'crucial' importance.

My own personal view is very much in sympathy with this, though it retains what I believe to be a highly significant difference. I would agree that the actual physical construction of the proposed 'thinking device' is likely to be crucial, but not merely with regard to the appropriateness of such terms as 'conscious awareness'. I think that the physical (or biological nature of the 'device' is likely to be crucial also in determining the effectiveness of its very operation. I would contend that the evolutionary development, through natural selection, of the ability to think consciously indicates that consciousness is playing an active role and has provided an evolutionary advantage to those possessing it. For various reasons I find it hard to believe that conscious awareness is merely a concomitant of sufficiently complex modes of thinking - and it seems to me clear that consciousness is itself functional.

People behave quite differently while in the states of full conscious awareness than they do when not properly conscious of what they are doing (such as while day-dreaming, sleepwalking or in a state of epileptic 'fugue'). Moreover, if consciousness had no operational effect on behaviour then conscious beings would never voice their puzzlement about the conscious state, and would behave just like unconcious mechanisms 'untroubled' by such irrelevancies! Moreover, I am very much aware that, in my own thinking, there are certain kind of problem for which conscious appraisal is essential,

whereas there are others - those which are 'automatic', there being a clearly defined algorithm which I have previously learnt (like walking or driving a car in very familiar circumstances) - for which conscious action is not needed.

Why do some kinds of thinking seem to 'need' consciousness while others seem not to? What is the crucial distinction between tasks that require (or at least benefit from) conscious thought and those that appear to have no real use for it? I do not suppose that very clear answers to these questions can be given at present. But at least there seems to be something very suggestive about the kind of terminology that most readily springs to mind. Terms such as 'automatic', 'following rules mindlessly', 'programmed' and even 'algorithm' seem appropriate for various kind of mental processes that are not necessarily accompanied by conscious awareness. These are the kind of activities that most closely accord with the actions of computers as we presently understand them. On the other hand, 'common sense', 'judgement of truth', 'understanding' and 'artistic appraisal' are among the terms one might use to describe mental activities that can seemingly be carried out effectively only when consciousness is present. For these, one is much more in the dark in attempting to imitate human thinking with computers. It seems not unnatural to suggest, therefore, that unconscious mental processes (perhaps such as can be carried out by the cerebellum) may indeed be effectively imitated by computers which do not differ in principle from those already in existence, but that conscious thinking may be essentially 'non-algorithmic' in its nature and qualitatively different from that which can

in principle be carried out by such computers.

The matter is perhaps not so clear-cut as this, however. In the first place it is evident that many unconscious ingredients are present in our artistic appraisals, in our strudied judgements and in our common sense. Though we form our judgements with consciousness as an apparently necessary ingredient, we do so by bringing together innumerable factors of logic, experience and prejudice, a good proportion of which we are probably not aware. Perhaps conscious appraisals are no less algorithmic in nature than unconscious 'automatic' actions, but are simple incomparably more complicated and subtle in their operation - as many would argue. Perhaps, then, consciousness is simply a concomitant of these higher level and more sophisticated - but still algorithmic - operations of mental machinery.

This is a view that gains support from the kind of picture that is now most commonly presented for the essential workings of the brain. A complicated network of neurones and synapses is envisaged, where electro-chemical signals propagate, indulging in elementary 'logical' switching processes - apparently similar to those of an electronic computer, though with a probabilistic ingredient. This all seems to be something that could be imitated closely by a computer - at least in principle. The additional probabilistic ingredient would seem to present no real problem. Computers can, after all, easily generate sequences of numbers which to all intents and purposes behave as random sequences. Or, if preferred, (effectively) random sequences can be produced by suitable physical processes (e.g.

decays of radio-active atoms) which could easily be additionally incorporated into the computer.

In any case - apart from the fact that the storage capacity is finite, rather than being unlimited - is one not simply providing an instance of a Turing machine? Any computational procedure whatever (in the normal sense of that term) comes under the heading of a procedure attanable by a Turing machine. To perform a computation, one is using a piece of physical apparatus, so the common argument goes, be it a human brain or an electronic computer or whatever, which is subject to the normal physical laws. These laws are of the usual deterministic type (classical, or else Schrödinger, time-evolution) or entirely probabilistic (state-vector reduction) and, as it would seem at first sight, are of the type which could be carried out by a Turing machine (with a limited store) with, if need be, an additional random input.

However the question of conputability, with ordinary deterministic laws, is not quite so straight-forward as it might seem. Indeed, according to Pour-El and Richards (1981, 1983), there is one sense in which evolving even so simple a deterministic system as the wave equation in flat space can be regarded as a non-computable procedure. But this particular sense is probably not the appropriate one in the present context. As they point out, sufficiently smooth solutions of the wave equation are indeed quite adequately computable from initial data (say by means of the Kirchhoff formula) and the same may be basically true also of the other main deterministic equations of physics - even though there will be the problems of stability, etc., that I mentioned in §1.

On the other hand, there is a critical assumption built into this discussion that the operative physical laws are indeed of the type referred to, where deterministic classical (or Schrödinger) evolution is maintained, except where punctuated by the random procedure of state-vector reduction. As I have been arguing in the earlier sections, one needs some replacement for this unhappy combination, the evolution and reduction being seen as approximations to some more accurate and precise grand scheme. It seems to me to be quite on the cards that this sought-for grand scheme contains an effectively non-computable time-evolution. It would be my view that the brain in its conscious state (or the 'mind'?) is able to harness these effects and make useof whatever subtleties are involved at that delicate borderline between linear Schrödinger evolution and the apparent randomness of reduction, thereby achieving effects far beyond those attainable by the ordinary operations of algorithmic computers.

Non-computability of some kind seems also to be a necessary ingredient of strong determinism. If one could compute the full space-time state of the universe, one could in principle compute one's own future. Then, of course, one might simply choose to do something different - an incompatibility with the feeling of free will.

These ideas are, of course, matters of considerable speculation, as present physical and physiological understandings go. Yet there are very many instances where biological systems have been able to make extraordinarily effective use of subtle physical or chemical effects. Surely the greatest of evolution's

achievements, the development of conscious awareness, must have needed to call upon physical processes even more profound and delicate than any we have yet understood. Any world in which minds can exist must be organized on principles far more subtle and beautifully controlled than those even of the magnificent physical laws that have so far been uncovered. At least, that is my own very strong opinion.

References

Born, M. (1927) Nature, 119, 354

Bohm, D.J. (1980) Wholeness and the Implicate Order (Routledge and Kegan Paul)

Deutsch, D. (1985) Three connections between Everett's interpretation and experiment, in Quantum Concepts in Space and Time, eds. C.J. Isham and R. Penrose (Oxford Univ. Press, Oxford)

DeWitt, B.S. and Graham, R.D. eds. (1973) The Many-Worlds Interpretation of Quantum Mechanics (Princeton Univ. Press, Princeton, N.J.)

Everett, H. (1957) "Relative state" formulation of quantum mechanics Revs. Mod. Phys. 29, 454-62

Hawking, S.W. and Ellis, G.F.R. (1973) The Large Scale Structure of Space-Time (Cambridge Univ. Press, Cambridge)

Feynman, R.P., Leighton, R.B. and Sands M. (1965) The Feynman Lectures on Physics, Vol. III (Addison Wesley, Reading Massachusetts)

Hofstadter, D.R. (1979) Godel, Escher, Bach: an External Golden Braid (Harvester Press Ltd., Hassocks, Sussex)

Karolyhazy, F. (1966) Nuovo Cimento A42, 390

Karolyhazy, F., Frenkel, A. and Lukacs, B. (1985) On the possible role of gravity in the reduction of the wave function, in Quantum Concepts in Space and Time, eds. C.J. Isham and R. Penrose (Oxford Univ. Press, Oxford)

Komar, A.B. (1969) Int. J. Theor. Phys. 2, 157

Komar, A.B. (1971) Semantic Foundation of the Quantization Program, in Studies in the Foundations, Methodology and Philosophy of Science, Vol. 4, Problems in the Foundations of Physics, ed. M. Bunge (Springer-Verlag, Berlin)

Ott, E. (1981) Strange attractors and chaotic motions of dynamical systems, Revs. Mod. Phys. 53, 655-671

Pais, A. (1982) Subtle is the Lord ... (Oxford Univ. Press, Oxford, New York)

Pearle, P. (1976) Phys. Rev. D13, 857

Pearle, P. (1984) Phys. Rev. D29, 235

Pearle, P. (1985) Models for Reduction, in Quantum Concepts in Space and Time, eds. C.J. Isham and R. Penrose (Oxford Univ. Press, Oxford)

Penrose, R. (1965) Revs. Mod. Phys. 37, 215-220

Penrose, R. (1979) Singularities and time-asymmetry, in General Relativity - An Einstein Centenary Survey, eds. S.W. Hawking and W. Israel (Cambridge Univ. Press, Cambridge)

Penrose, R. (1981) Time-asymmetry and quantum gravity, in Quantum Gravity 2 (Oxford Univ. Press, Oxford)

Penrose, R. (1985) Quantum gravity and state-vector reduction, in Quantum Concepts in Space and Time (Oxford Univ. Press, Oxford)

Penrose, R. and Rindler, W. (1985) Spinors and Space-Time, Vol. 2; Spinor and Twistor Methods in Space-Time Geometry (Cambridge Univ. Press, Cambridge)

Pour-El, M.B. and Richards, I. (1981) Adv. in Math. 39, 215-239

Pour-El, M.B. and Richards, I. (1983) Adv. in Math. 48, 44-74

Searle, (19)
in The Mind's I, eds.

Turing, M. (1950) Computing Machinery and Intelligence, Mind, 59; also in Minds and Machines, ed. A.T. Anderson (McGraw Hill, New York 1978)

von Neumann, J. (1955) Mathematical Foundations of Quantum Mechanics (Princeton Univ. Press, Princeton)

Wald, R. (1984) General Relativity (Chicago Univ. Press, Chicago)

Wigner, E.P. (1961) Remarks on the mind-body question, in The Scientist Speculates ed. I.J. Good (Heinemann, New York); also in Quantum Theory and Measurement eds. J.A. Wheeler and H. Zurek (Princeton Univ. Press, Princeton, N.J. 9183)

171
LOCAL TWISTOR TRANSPORT AT $\mathscr{I}^+$: AN APPROACH TO THE GOOGLY

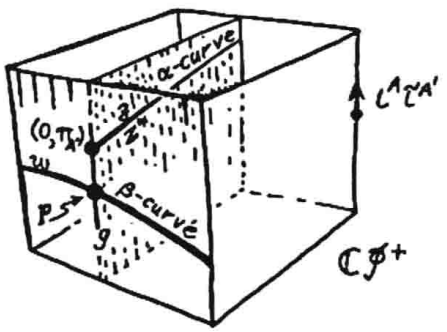
Twistor Newsletter **23** (1987). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 Local twistor transport at $\mathscr{I}^+$: An approach to the googly. *Twistor Newsletter* **23** (1987). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.17 Local Twistor Transport at $\mathscr{I}^+$: an Approach to the Googly *by R. Penrose* **(TN 23, February 1987)**

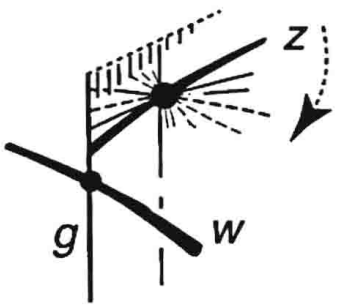
This note reports on some work in progress concerning the googly graviton, and suggests a certain shift from the viewpoint I have described in §III.4.11, §III.4.12 and §III.4.15. This shift results from an attempt to relate the structure of the googly twistor space

to local twistor transport on $\mathbb{C}\mathscr{I}^+$. (Several of the ideas described here have resulted from discussions with Vanessa Thomas.)



The central idea is to use local twistor transport along β-curves on $\mathbb{C}\mathscr{I}^+$ to provide a family of 'flat coordinate systems' for $\mathbf{T}$ ($\mathbf{T}$ = the non-projective blown-up twistor space, which for flat space twistors is the space of $\dot{Z}Z\rceil = Z^\alpha Z^\beta I_{\beta\gamma}$). Recall that apart from generators of $\mathbb{C}\mathscr{I}^+$, the *null geodesics* on $\mathbb{C}\mathscr{I}^+$ are of two kinds: *β-curves* (lying on α-planes on $\mathbb{C}\mathscr{I}^+$ and representing projective asymptotic *dual twistors*) and *α-curves* (lying on β-planes on $\mathbb{C}\mathscr{I}^+$ and representing projective asymptotic *twistors*). Let w be a β-curve. Then local twistor transport along w can be used to establish an identification between all the local twistor spaces at the various points of w. Let us call the resulting (flat) identified twistor space $\mathbb{T}(w)$. Now any α-curve $\mathfrak{z}$ which does not lie too far from the non-singular region of w will meet a unique generator g of $\mathbb{C}\mathscr{I}^+$ intersecting w. Assigning an appropriate scaling to $\mathfrak{z}$ (given by a $\pi_{A'}$ spinor parallelly propagated along $\mathfrak{z}$, for which $\iota_A\pi_{A'}$ is a tangent covector to $\mathfrak{z}$, the tangents to generators of $\mathbb{C}\mathscr{I}^+$ being $\iota^A\bar{\iota}^{A'}$) we get an asymptotic twistor Z^α. With respect to any point of $\mathfrak{z}$, Z^α has the local twistor description $(0,\pi_{A'})$. This can be carried along g to the point $p = g \cap w$ by local twistor transport, and Z^α is thereby labeled by a unique element of $\mathbb{T}(w)$. Thus, the elements of $\mathbb{T}(w)$ provide 'coordinates', locally, for the *asymptotic twistor space* $\mathscr{T}$. For points of $\mathscr{T}$ which do not lie on I ($\subset \mathscr{T}$), these provide smooth coordinates. However, the coordinates will not generally be smooth at points of I. The 'blown-up' local twistors $\dot{Z}Z\rceil$ ($\in \mathbb{T}^{(\alpha\beta)}(w)I_{\beta\gamma}$) have a better chance of being smooth 'coordinates' locally at points of $\mathbb{I} \subset \mathscr{T}$, where $\mathscr{T}$ is the *blown-up* (non-projective) twistor space $\mathbb{T}$. I do not yet know if — or under what circumstances — they *are* smooth. My main doubts would arise when $\dot{Z}$ approaches I by having $\mathfrak{z}$ swing to the vertical, about some

point *not* on g:



We need to be careful about the interpretation of $\dot{Z}$ in terms of $\mathbb{T}(w)$. The infinity twistor Π $(I_{\alpha\beta})$ is well-defined as a local twistor and is local-twistor *constant* over the whole of $\mathbb{C}\mathscr{I}^+$. (Indeed, it is constant, in any *vacuum* space-time, throughout the *finite* regions also.)

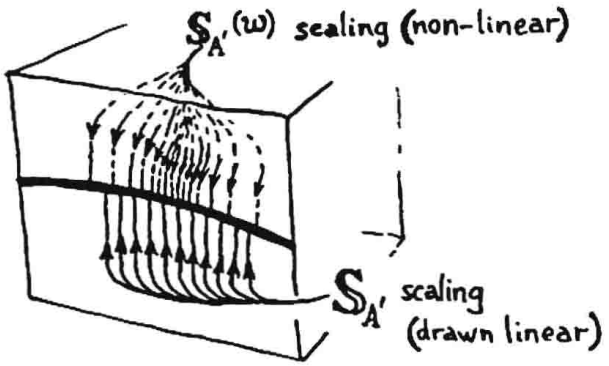
$$I_{\alpha\beta} = \begin{bmatrix} 0 & -\mathrm{i}\,\iota_A\tilde{\iota}^{B'} \\ \mathrm{i}\,\tilde{\iota}^{A'}\iota_B & 0 \end{bmatrix} \quad \text{at } \mathbb{C}\mathscr{I}^+,\ \rho' = 0$$

The form that Π takes does not depend upon the structure (news function, radiation field) of $\mathbb{C}\mathscr{I}^+$. However, the 'π-space' $\mathbb{S}_{A'}(w)$ of $\mathbb{T}(w)$ does not agree with the asymptotic 'π-space' $\mathbb{S}_{A'}$ that is defined for $\mathbb{C}\mathscr{I}^+$ as a *whole* (cf. Penrose & Rindler 1986, p. 411, 417; here $\mathbb{S}_{A'} = \mathbb{S}_{A'}(\mathcal{S})$) for any cut $\mathcal{S}$ of $\mathscr{I}^+$– in the sense of 2-surface twistors). The difference between $\mathbb{S}_{A'}(w)$ and $\mathbb{S}_{A'}$ was extensively discussed, in effect, by Burnett-Stuart (1982), the elements of $\mathbb{S}_{A'}$ being his 'π's and the elements of $\mathbb{S}_{A'}(w)$, his 'p's (cf. also §III.4.13), though he was not concerned there with local twistor transport. In Penrose and MacCallum (1973), the difference between local twistor transport of $\mathcal{Z}$ twistors ('twistors entirely on $\mathscr{I}$'), which gives $\mathbb{S}_{A'}(w)$, and a *modified* local twistor transport, which gives $\mathbb{S}_{A'}$ was pointed out (p. 301). The modification results from suppressing some of the curvature terms in $P_{ABA'B'}$ which appear in the local-twistor propagation of $\pi_{A'}$s. The suppressed terms are those of the form $P_{11A'B'}$ and $P_{AB1'1'}$, these being essentially the anti-self-dual and self-dual parts, N and $\tilde{N}$, of the *news function*, respectively:

$$P_{11A'B'} = N\tilde{\iota}_{A'}\tilde{\iota}_{B'}, \quad P_{AB1'1'} = \tilde{N}\iota_A\iota_B.$$

The relation between $\mathbb{S}_{A'}(w)$ and $\mathbb{S}_{A'}$ cannot extend globally unless this relation is linear. The presence of self-dual news function $\tilde{N} \neq 0$ implies that the relation is

nonlinear. It is not global because w is not a $\mathbb{CP}^1$, but enters the singular region of $\mathbb{C}\mathscr{I}^+$ (where $\tilde{N}$ blows up) and so is non-compact. (It is only $\tilde{N}$, and not N, which enters into the local twistor propagation along w. N would show up, on the other hand, if we were to propagate along α-curves.) In the figure below I have schematically indicated the nonlinear relation between the $\mathbb{S}_{A'}(w)$ scalings and the $\mathbb{S}_{A'}$ scalings.

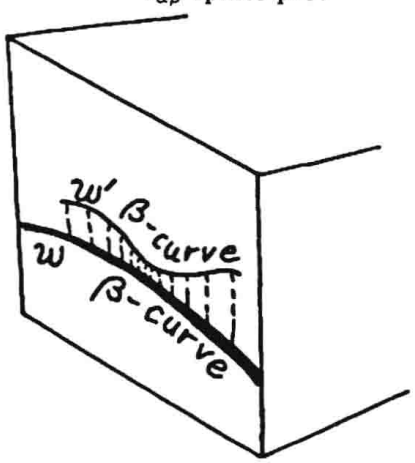


In Penrose & Rindler (1986) (pp. 374–377), a connection between the strong conformal geometry of $\mathscr{I}^+$ (angle and null-angle structure) and the product $\sqcup\!\sqcap$ of the infinity twistors was pointed out. In effect, we can interpret $\sqcap$ as providing a structure which converts tangent vectors in a β-plane to scalings up the generators. More correctly, it converts spinors λ^A (take $\lambda^A\tilde{\iota}^{A'}$ to be tangent to w) to multiples of $\tilde{\iota}^{A'}$:

$$\lambda^A \;\mapsto\; \lambda^A(-\mathrm{i}\iota_A\tilde{\iota}^{B'}) \;=\; (-\mathrm{i}\lambda^A\iota_A)\tilde{\iota}^{B'}.$$

$\uparrow$

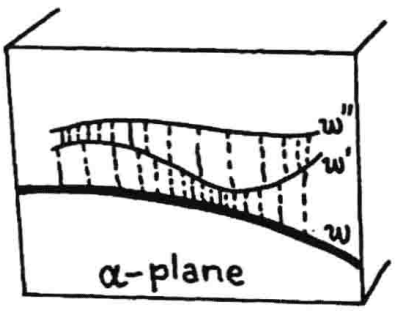
$I_{\alpha\beta}$ spinor part



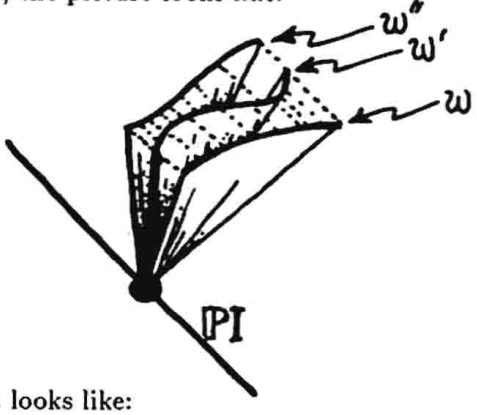
(To achieve the conversion of tangent vectors, we need a fixed scaling in the $\nearrow$ direction in the fig.). The scaling up the generators allows us to locate the 'next' β-curve w'.

 Local twistor transport at $\mathscr{I}^+$: An approach to the googly. *Twistor Newsletter* **23** (1987). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

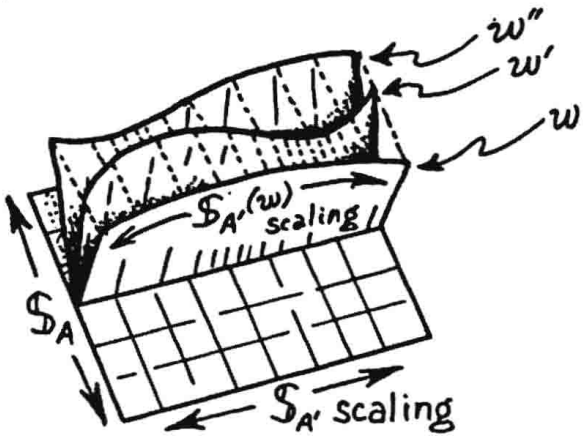
Actually there is a family of such scalings, related by bilinear transformations in the Riemann sphere of $\mathbb{S}_{A'}(w)$. These give all the various possible 'next' β-curves. The *curvature* $\tilde{\Psi}_4$ (self-dual gravitational radiation field) of each α-plane on $\mathbb{C}\mathscr{I}^+$ is measured by the derivative of $\tilde{N}$ up the generators, and it gives the relation between w, a 'next' β-curve w' and a 'next-but-one' β-curve w''.



The *googly* structure of asymptotic twistor space must find a way of encoding this curvature. In terms of $\mathbb{P}\mathscr{T}$, the picture looks like:



whereas in terms of $\mathbb{P}\mathscr{T}$, it looks like:



It would seem that in order to incorporate the *local* structure afforded by $\sqcap$, which allows the $\mathbb{S}_{A'}(w)$ spaces to be described, rather than simply the global $\mathbb{S}_{A'}$ structure, we need something like the form $\widehat{Z dZ}$ (which is local) rather than $Z\rceil$ (which, as a structure on $\mathscr{T}$ or $\mathscr{T}$, requires the global space $\mathbb{S}_{A'}$). Recall that in §III.4.11, §III.4.12 and §III.4.15 I considered the forms $Z\rceil Z\rceil \mathscr{D}Z$, $Z\rceil Z\rceil d^4Z$ and things like $f_{-6}Z\rceil Z\rceil \mathscr{D}Z$, where $\mathscr{D}Z = \frac{1}{6}\, dZ \wedge dZ \wedge dZ \wedge dZ$ and $d^4Z = \frac{1}{24}\, dZ \wedge dZ \wedge dZ \wedge dZ$, the point being that $Z\rceil Z\rceil \mathscr{D}Z$, and $Z\rceil Z\rceil d^4Z$ are finite on $\mathbb{I}$, whereas $\mathscr{D}Z$ and d^4Z are not, and moreover $d(fZ\rceil Z\rceil \mathscr{D}Z) = 0$ iff f has homogeneity degree -6. The present point of view is that we need the more local structure afforded by $\widehat{ZdZ}$, or say $d\widehat{Z\wedge dZ}$, in order to allow $\mathbb{S}_{A'}(w)$ to appear naturally rather than just $\mathbb{S}_{A'}$. It is presumably significant, therefore, that the 'form'

$$\sqcap Z = \widehat{ZdZ} \otimes \mathscr{D}Z$$

(which has Young symmetry $\begin{smallmatrix}\square\\\square\end{smallmatrix}$, the skew part 'showing') is also finite on $\mathbb{I}$, and so is $\widehat{ZdZ}\otimes d^4Z$ (Young symmetry $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$), whereas $d\widehat{Z\wedge dZ}\otimes \mathscr{D}Z$ and $d\widehat{Z\wedge dZ}\otimes d^4Z$ are not. Note that $\sqcap Z$ is 'almost' the blown-up twistor $\dot{Z}Z\rceil$ *itself* — or, more exactly $\widehat{\dot{Z}Z\rceil}\,\|\|$ — where we snuff off the 'legs' with $d\dot{Z}$'s. I do not yet adequately see how to incorporate -6 functions in a natural way. Perhaps one needs some sort of invariant derivative which can act on $f_{-6}\sqcap Z$. (Simply 'd' won't do, since it doesn't invariantly act on $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ things. We got away with it before — more or less — because the 'legs' of $Z\rceil Z\rceil \mathscr{D}Z$ were treated as constant, though this may well be inappropriate.) Note that $f\sqcap Z$ is Lie constant along Υ (Euler field) iff f is deg. -6.

To conclude, let us consider some relevant transformations of a flat twistor space $\mathbb{T}$. These correspond to: (a) supertranslations of $\mathbb{C}\mathscr{I}^+$ which are self-dual (or should I call them anti-self-dual?) translations in the sense that they preserve the linear (or affine) structure of the β-planes on $\mathbb{C}\mathscr{I}^+$, and so preserve α-curves (whence they can act directly on $\mathbb{T}$); these are given by

$$\dot{Z} \mapsto \dot{Z} + \text{❡}$$

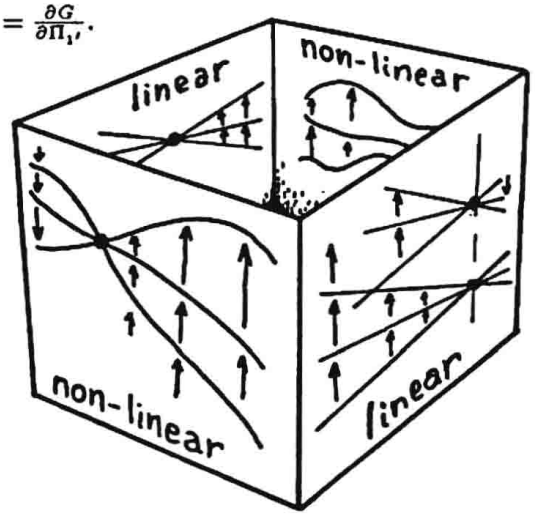
where ❡ is a function of $Z\rceil$ homogeneous of degree 1 — and (b) transformations of $Z\rceil$ which send $\mathbb{S}_{A'}$ to things like $\mathbb{S}_{A'}(w)$ above, when $\tilde{N} \neq 0$; these are given by transformations of $\pi_{A'}$ preserving $\delta\pi = \pi_{A'}d\pi^{A'} = \widehat{ZdZ}$. The transformations (a) take the form

$$\begin{cases} \omega^A \mapsto \omega^A + \Omega^A(\pi_{A'}) \\ \pi_{A'} \mapsto \pi_{A'} \quad \uparrow \text{homogeneous deg. 1} \end{cases}$$

whereas (b) are [generated by (a) and] those of the form

$$\begin{cases} \omega^A \mapsto \omega^A \\ \pi_{A'} \mapsto \Pi_{A'}(\pi_{B'}) \end{cases}$$

where $\Pi_{A'}$ can be given by a generating function $G(\pi_{0'}, \Pi_{1'})$ homogeneous of degree 2, with $\pi_{1'} = \frac{\partial G}{\partial \pi_{0'}}$, $\Pi_{0'} = \frac{\partial G}{\partial \Pi_{1'}}$.



It is noteworthy that *both* (a) and (b) *preserve both* the forms $\widehat{ZdZ}$ and $\mathscr{D}Z$, so they certainly preserve $\lceil Z$ (and $d\hat{Z}\wedge d\hat{Z}$ and d^4Z and Υ).

Further work is in progress, see §III.4.1.

References

Burnett-Stuart, G. (1982) D.Phil. thesis, Oxford University.

Penrose, R. & MacCallum M.A.H. (1973) Twistor theory: an approach to the quantisation of fields and space-time. Phys. Rep. **6C**, 241–316.

Penrose, R. & Rindler, W. (1986) *Spinors and Space-Time*, vol. 2, Cambridge University Press, Cambridge.

172
AN APPROACH TO A COORDINATE-FREE CALCULUS AT $\mathbb{I}$

Twistor Newsletter **24** (1987). With V. Thomas. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

Vanessa Thomas became my wife on 9 April 1988.

 An approach to a coordinate-free calculus at $\mathbb{I}$. *Twistor Newsletter* **24** (1987). With V. Thomas. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.18 **An approach to a coordinate free calculus at $\mathbb{I}$** *by R. Penrose and V. Thomas* (TN 24, September 1987)

In attempting to obtain an effective googly graviton construction, it appears that it would be helpful to have a way of calculating at (and near) $\mathbb{I}$ (the non-projective blow-up of I in blown-up twistor space $\mathbb{T}$). $\mathbb{T}$ is defined by "coordinates" $(\omega^A \pi_{B'}, \pi_{A'} \pi_{B'})$ — i.e., by $\dot{Z}\, \hat{Z}$ — and $\mathbb{I}$ arises in the limit $\lambda \to 0$ of $(\lambda^{-1}\omega^A, \lambda \pi_{A'})$. Various calculations have been carried out earlier (by Roger Penrose, Toby Bailey and Michael Singer) but only using particular coordinate systems. What we need is an invariant way of seeing when a tensor quantity in $\mathbb{T}$ is finite at $\mathbb{I}$, etc.

A set of local coordinates for $\mathbb{T}$ is provided by

$$w = (\pi_{0'})^2, \quad x = \omega^0 \pi_{0'}, \quad y = \omega^1 \pi_{0'}, \quad v = -\omega^1 \pi_{1'}.$$

(Other similar sets of coordinates would be needed to cover other parts of $\mathbb{T}$.) Here $w = 0$ gives $\mathbb{I}$, so we are interested in how quantities behave as $w \to 0$. Taking a twistor basis $-\circ\!\sqcup$, $\circ\!\sqcup$, $\circ\!\downarrow$, $\bullet\!\downarrow$ and dual basis $\circ\!\uparrow$, $\bullet\!\uparrow$, $\bullet\!\sqcap$, $-\circ\!\sqcap$ where $\circ\!\uparrow\bullet\!\sqcup = 1 = \bullet\!\sqcap\bullet\!\downarrow$, $\circ\!\uparrow\circ\!\downarrow = \bullet\!\uparrow\bullet\!\downarrow = \bullet\!\uparrow\circ\!\downarrow = \bullet\!\uparrow\bullet\!\downarrow = 0$, we have $w = \hat{Z}\bullet\, \hat{Z}\bullet$, $x = \overset{\circ}{Z}\, \hat{Z}\bullet$, $y = \overset{\bullet}{Z}\, \hat{Z}\bullet$, $z = \overset{\bullet}{Z}\, \hat{Z}\circ$ whence $d\hat{Z}\bullet = \frac{1}{2}\dfrac{dw}{\hat{Z}\bullet}$, $d\hat{Z}\circ = \dfrac{dv}{\overset{\bullet}{Z}} - \dfrac{\hat{Z}\circ}{\overset{\bullet}{Z}}\left(\dfrac{dy}{\hat{Z}\bullet} - \dfrac{\overset{\bullet}{Z}\, dw}{2(\hat{Z}\bullet)^2}\right)$, $d\overset{\bullet}{Z} = \dfrac{dy}{\hat{Z}\bullet} - \dfrac{\overset{\bullet}{Z}\, dw}{2(\hat{Z}\bullet)^2}$, $d\overset{\circ}{Z} = \dfrac{dx}{\hat{Z}\bullet} - \dfrac{\overset{\circ}{Z}\, dw}{2(\hat{Z}\bullet)^2}$.

These can be substituted into a differential form of interest, and then $\circ\!\downarrow$, $\bullet\!\downarrow$, $\circ\!\uparrow$, $\bullet\!\uparrow$ eliminated by allowing them to vary. For example, one important differential form is

$$\hat{Z}\, d\, \hat{Z} = \pi_{0'} d\pi_{1'} - \pi_{1'} d\pi_{0'}.$$

We know from previous work that $\hat{Z}\, d\, \hat{Z}$ vanishes at $\mathbb{I}$. Can we see this in a coordinate-free way? First, let us *use* our coordinates: direct calculation from the above "blob" notation gives

$$\overset{\bullet}{Z}\, \hat{Z}\, d\, \hat{Z} = d(\overset{\bullet}{Z}\, \hat{Z})\, \hat{Z}\sqcup_{\circ\bullet}$$

which, by $\sqcap\!\sqcup = -\sqcap\sqcap$ gives

$$\overset{\bullet}{Z}\, \hat{Z}\, d\, \hat{Z} = -d(\overset{\bullet}{Z}\, \hat{Z})\, \hat{Z}.$$

We can now *remove* $\overset{\bullet}{\downarrow}$, since this basis vector can be chosen arbitrarily, so multiplying by $\overline{Z}|$ we get the identity

$$\overset{\downarrow}{Z}\,\overline{Z}|\overline{Z\,d\overset{\downarrow}{Z}} = \overline{Z}|\overline{Z\,d(\overset{\downarrow}{Z}\,\overset{\downarrow}{Z})}$$

(which we can now see directly — but we might not have guessed this expression without first using $\overset{\bullet}{\downarrow}$, $\overset{\circ}{\downarrow}$ ...). This identity indeed tells us that $\overline{Z\,d\overset{\downarrow}{Z}}$ is $\mathcal{O}(w)$, i.e., $\mathcal{O}(\overline{Z}|\,\overline{Z}|)$.

The $\sqcap$ of $\overline{Z\,d(\overset{\downarrow}{Z}\,\overset{\downarrow}{Z})}$ does double duty — i.e., making *both* the $\overset{\downarrow}{Z}$ and the $d(\overset{\downarrow}{Z})$ hooked (since it can be rewritten $\sqcap| = -\sqcap\!\!\sqcap|$) — telling us that the R.H.S. is $\mathcal{O}(\overline{Z}|\overline{Z}|)$, being of the form $\overline{Z}|\,\overline{Z}|\times d\,(\overset{\downarrow}{Z}\,\overline{Z})|$, while the L.H.S = non-singular non-zero $((\overset{\downarrow}{Z}\,\overline{Z})|)$ × required form $\overline{Z\,d\overset{\downarrow}{Z}}$. Therefore $\overline{Z\,d\overset{\downarrow}{Z}} = \mathcal{O}(\overline{Z}|\overline{Z}|)$.

Now consider the 3-form $\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}$. Calculating in terms of the above coordinates, and eliminating the basis elements, as before, we get

$$\overset{\downarrow\!\!\downarrow}{Z}\,\overline{Z}|\overline{Z}|\overline{Z}|\;\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}} = \frac{1}{2}\,\overline{Z}|\underline{\overline{Z}|d(\overset{\downarrow}{Z}\,\overline{Z})|}\wedge d(\overset{\downarrow}{Z}\,\overline{Z})\wedge d(\overset{\downarrow}{Z}\,\overline{Z})$$

$$+\frac{1}{3}\,d(\overset{\downarrow}{Z}\,\overline{Z})|\wedge\underline{d(\overset{\downarrow\!\!\downarrow}{Z}\,\overline{Z})\wedge d(\overline{Z}|\overline{Z}|)}$$

$$-\frac{1}{6}\,\overset{\downarrow}{Z}\,\overline{Z}|\underline{d(\overline{Z}|\overline{Z}|)}\wedge d(\overset{\downarrow}{Z}\,\overline{Z})\wedge d(\overset{\downarrow}{Z}\,\overline{Z})$$

$$\propto \left|\overset{\downarrow}{Z}\,\overset{\downarrow}{Z}\,\overset{\downarrow}{Z}\,\overline{Z}\;\overline{Z\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}\right|$$

The left-hand side being $\mathcal{O}(\overline{Z}|\,\overline{Z}|)\times\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}$ and the right-hand side being $\mathcal{O}(1)$ (since $|\overset{\downarrow}{Z}\,\overline{Z}$ and $|d(\overset{\downarrow}{Z}\,\overline{Z}) = |d\overset{\downarrow}{Z}\,\overline{Z}$ are both $\mathcal{O}(1)$). Thus $\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}$ has a simple pole at $\mathbb{I}$, so $\overline{Z}|\,\overline{Z}|\,\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}$ and $\overline{Z\,d\overset{\downarrow}{Z}}\propto\overline{\overset{\downarrow}{Z}\,d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}\wedge d\overset{\downarrow}{Z}}$ are both finite at $\mathbb{I}$ — as is required for various googly ideas (Penrose §III.4.11, §III.4.12 and §III.4.17).

As for vector fields, we can find that, for any $\overset{\downarrow}{\nabla}$,

$$\overset{\downarrow}{\nabla}_Z = \overset{\downarrow}{\nabla}\overset{\downarrow}{Z}\,\frac{\partial}{\partial(\overline{Z}\,\overline{Z})} + \overset{\circ}{\nabla}\overset{\downarrow}{Z}\,\frac{\partial}{\partial(\overset{\circ}{Z}\,\overline{Z})} + \overset{\bullet}{\nabla}\overset{\downarrow}{Z}\,\frac{\partial}{\partial(\overset{\bullet}{Z}\,\overline{Z})} + \overset{\bullet}{\nabla}\overset{\downarrow}{Z}\,\frac{\partial}{\partial(\overset{\bullet}{Z}\,\overline{Z})}.$$

The coefficients are all components of $\overset{\downarrow}{\nabla}\overset{\downarrow}{Z}$, so we see that

$\overset{\downarrow}{\nabla}$ is well-behaved at $\mathbb{I}$ *if* $\overset{\downarrow}{\nabla}\overset{\downarrow}{Z}$ is non-singular

612 An approach to a coordinate-free calculus at 𝕀. *Twistor Newsletter* **24** (1987). With V. Thomas. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

In particular, we can check that the vector fields representing "googly maps" are appropriately well behaved, using this procedure.

By combining this with the above procedure for forms, we should be able to deal with any tensor field $\mathbf{T}$.

173
TWISTORS IN GENERAL RELATIVITY

General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986). Ed. M. A. H. MacCallum, 158–76. Cambridge University Press, 1987

 Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

Twistors in General Relativity

by
Roger Penrose
Mathematical Institute
24-29 St. Giles
Oxford

Abstract

The two most significant definitions of a twistor in (conformally) flat space-time are that given by solutions of the twistor equation and that given by α-planes in the complexification of the space-time. It is shown that these concepts are naturally dual to one another, and that they give rise to various different twistor concepts in curved space-time when referred to submanifolds of differing dimension. One of these concepts, that of 2-surface twistors, gives rise to a quasi-local definition of mass (and angular momentum). The status of this definition is explored in relation to various examples.

1. Introduction

There are many different ways of introducing the ideas of twistor theory. Fortunately there are now two reasonably comprehensive books *available on the subject (Penrose and Rindler 1986, Huggett and Tod* 1985) and it will not be necessary for me to be in any way complete in my descriptions of twistors.

The most primitive idea of a twistor can be taken to be a (geometrical) light ray in space-time, i.e. a null geodesic. In twistor theory one transfers one's attention from the space-time to twistor space (the space of twistors); so according to this primitive picture, we think of a space-time point in terms of the family - a (Riemann) sphere's worth - of light rays which pass through that point. Thus, a space-time point is interpreted, twistorially, as a Riemann sphere in twistor space, whereas a light ray in space-time is interpreted as a single point in twistor space.

There are various motivations which underlie this procedure. Quantum gravity could allow the space-time points actually to become "fuzzy" - rather than merely the light cones in space-time becoming fuzzy - this being allowed for if it is the twistor space, rather than the space-time which is taken as primary. Moreover, there is the possibility that one might be able to find an alternative to a space-time description at curvature singularities, so that some kind of coherent physical description could still persist at the big bang, or at black-hole cores,

Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

even when our conventional notions of space-time give out. Twistor descriptions are essentially non-local, which could fit in with the known non-local features of quantum theory (as manifested, for example, in the EPR-Bohm type of experiment of Freedman and Clauser 1972, Aspect, Grangier and Roger 1982). Twistors have an essential chirality, which fits in with the known chiral behaviour of Nature as manifested in weak interactions. Massless fields and conformally invariant interactions seem to have a primitive role to play in Nature, these being very naturally treated within twistor theory. Perhaps most importantly of all, twistors are fundamentally complex (holomorphic) objects and serve to unify the basic complex numbers of quantum theory (probability amplitudes) with the very structure of space-time.

2. Twistors in Flat Space-time

I shall concentrate on just two of the various different more complete ways of regarding a twistor in standard (four-dimensional, signature +---) Minkowski space-time $\mathbb{M}$. The first is as a solution of the twistor equation

$$\nabla^{(A}_{A'}\omega^{B)} = 0.$$

(See Penrose and Rindler 1984 for the conventions concerning the 2-spinor notation used here.) The solution space of this equation is four-complex-dimensional, and we have

$$\omega^A = \mathring{\omega}^A - ix^{AA'}\mathring{\pi}_{A'},$$

$$\pi_{A'} = \mathring{\pi}_{A'},$$

where $x^a (= x^{AA'})$ is the position vector relative to some fixed origin O, and where the symbols with the small circle above denote constant spinors, whose values coincide with that of the corresponding spinor field (denoted without the circle above) at the origin O. The pair of spinor fields ω^A and $\pi_{A'}$ are the spinor parts of the twistor Z^α (where α is a four-dimensional abstract index), and we may write

$$Z^\alpha = (\omega^A, \pi_{A'}).$$

The second way of viewing a twistor that I am considering here, is in terms of an α-plane. For this we need the complexification $\mathbb{C}\mathbb{M}$ of Minkowski space $\mathbb{M}$ - and , moreover, the conformal compactification $\mathbb{C}\mathbb{M}^\#$ of this. An α-plane is a complex 2-plane which is totally null, in the sense that the (complexification of the) space-time metric g_{ab} vanishes

identically over the entire plane. In fact there are <u>two</u> disjoint families of such totally null planes, namely the α-planes and the β-planes. Any complex null vector has the spinor form

$$\lambda^A \pi^{A'},$$

and the tangents to an α-plane are the vectors of this form for which $\pi^{A'}$ is held fixed and λ^A varies. Correspondingly the tangents to a β-plane have this form where λ^A is held fixed and $\pi^{A'}$ varies. Clearly all the tangent vectors to an α-plane (or β-plane) are orthogonal to one another (e.g. $(\lambda^A \pi^{A'})(\lambda_A \pi_{A'}) = 0$, since $\pi^{A'} \pi_{A'} = 0$), whence the induced metric vanishes identically. An α-plane (or a β-plane) which contains a real point, contains precisely a single light ray of real points (given by $\lambda^A = k\bar{\pi}^A$, $k \in \mathbb{R}$), and this regains our original <u>light-ray</u> way of viewing certain twistors - these being the <u>null</u> twistors.

The fact that α-planes in $\mathbb{C}\mathbb{M}$ have totally null tangent 2-planes of this type immediately implies that they are flat (totally geodesic) - by a Lie bracket argument. In fact we can take $\pi_{A'}$ to be <u>constant</u> over the α-plane (the "covariant" version $\pi_{A'}$ being preferred over the "contravariant" version $\pi^{A'}$ for the reason that this constancy is thereby rendered conformally invariant). This gives us an alternative definition of a (non-zero) twistor, as an α-plane with constant $\pi_{A'}$ associated with it. The freedom in the choice of $\pi_{A'}$ is simply that of a rescaling

$$\pi_{A'} \mapsto \chi \, \pi_{A'} \qquad (0 \neq \chi \in \mathbb{C}).$$

The choice of scaling for $\pi_{A'}$ fixes the <u>scaling for the twistor</u>.

The equivalence (in $\mathbb{C}\mathbb{M}^*$) between this definition and that provided by solutions of the twistor equation is seen from the fact that the <u>zero-locus</u> of any (non-zero) solution is an α-plane. The points with position vectors x^a, at which $\omega^A = 0$, are given by

$$\mathring{\omega}^A = ix^{AA'} \mathring{\pi}_{A'}.$$

If we start from one such point at which this equation holds, the directions along which we can move, while maintaining satisfaction of this equation, clearly have the form

$$\lambda^A \mathring{\pi}^{A'}$$

for arbitrary λ^A (and recall that $\mathring{\pi}^{A'} = \pi^{A'}$), showing that this locus is indeed an α-plane and the scaling is provided by $\pi_{A'}$ as before. We should note that the <u>constant</u> solutions of the twistor equation evade this description, since now $\pi_{A'} = 0$. For these, we invoke the conformal

invariance (that we shall examine shortly) of the twistor equation. By transforming null infinity to a finite place we observe that there is indeed a zero-locus of ω^A, which is an α-plane lying entirely at null infinity.

3. Duality Between the Two Definitions

I want to point out an essential duality that underlies the relation between these two definitions of a twistor. In fact dual twistors are identified with complex conjugate twistors, so it is the twistor concept associated with a β-plane which must be seen to be dual to that associated with a solution of the twistor equation. Indeed, we can use this duality to see what the significance of that particular equation that we have called the twistor equation actually is!

Suppose we simply have some spinor field ω^A which is taken to define some object that is naturally dual to the β-plane type of twistor, i.e. to totally null complex 2-planes (β-planes) in $\mathbb{CM}^\#$ with constant spinors λ_A attached to them, the tangent vectors to the 2-plane having the form

$$\lambda^A \xi^{A'}$$

for varying $\xi^{A'}$. What equation should ω^A satisfy? Now the constancy of λ_A over the β-plane asserts

$$\lambda^A \xi^{A'} \nabla_{AA'} \lambda_B = 0$$

for all $\xi^{A'}$, i.e.

$$\lambda^A \nabla_{AA'} \lambda_B = 0 \ .$$

The duality with the ω^A-field is expressed in the existence of a scalar product between the ω^A-field and the λ_A-scaled β-plane, which must be an element of $\mathbb{C}$. We take this scalar product to be

$$\omega^A \lambda_A ,$$

and the fact that it is a single element of $\mathbb{C}$ requires that this be constant over the β-plane:

$$\lambda^A \xi^{A'} \nabla_{AA'} (\omega^B \lambda_B) = 0$$

for all $\xi^{a'}$. Using the above relation we obtain $\lambda^A \lambda_B \nabla_{AA'} \omega^B = 0$, i.e.

$$\lambda_A \lambda_B (\nabla^{(A}_{A'} \omega^{B)}) = 0.$$

This must hold for all β-planes and thus for all λ_A, so the bracketed expression $\nabla^{(A}_{A'} \omega^{B)}$ must indeed vanish, which is the required twistor equation.

Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

4. The Geometry and Physics of Twistors

The α-plane and twistor equation concepts of a twistor lead naturally in somewhat different directions. First, the α-plane concept leads us directly into various geometrical considerations. We have seen that in the real space $\mathbb{M}$, the α-plane definition provides the picture of a (null) twistor as a (scaled) light ray. Also, in complex terms, the space $\mathbb{CM}^{\#}$ is seen to be the Klein representation of projective lines in projective twistor space $\mathbb{PT}$ (a complex projective 3-space). Points in $\mathbb{PT}$ correspond to β-planes in $\mathbb{CM}^{\#}$. Lines in $\mathbb{PT}$ correspond to points in $\mathbb{CM}^{\#}$. Pairs consisting of planes in $\mathbb{PT}$ with single points marked on them correspond to null geodesics in $\mathbb{CM}^{\#}$ - the intersections of α-planes with β-planes. The complex conjugate of an α-plane is a β-plane. Thus, in $\mathbb{PT}$, the complex conjugate of a point is a plane. For a null twistor, the point lies on the plane, and there is a corresponding real null geodesic in $\mathbb{M}^{\#}$.

The definition in terms of the twistor equation leads more directly to connections with physics. In the first place, any solution ω^A of the twistor equation acts as a spin-lowering operator for massless fields. If $\varphi_{AB\ldots L} = \varphi_{(AB\ldots L)}$, with n indices (i.e. spin n/2), satisfies the massless free-field equation

$$\nabla^A{}_{A'}\varphi_{AB\ldots L} = 0$$

then so does

$$\omega^L\varphi_{AB\ldots KL'},$$

with n-1 indices (i.e. spin (n-1)/2), since

$$\nabla^A{}_{A'}(\omega^L\varphi_{AB\ldots KL}) = \varphi_{AB\ldots KL}\nabla^{(A}{}_{A'}\omega^{L)} = 0.$$

Similarly ω^A can be used as a spin-raising operator since, by a direct calculation

$$\bar{\omega}^{M'}\nabla_{MM'}\varphi_{AB\ldots L} - i(n+1)\bar{\pi}_{(M}\varphi_{AB\ldots L)}$$

also satisfies the massless free-field equation.

The twistor equation plays yet another role: a twistor has an intepretation as providing the angular momentum $M^{ab}(= -M^{ba})$ and momentum p_a of a (classical) massless particle. The conditions that p_a be future-null and proportional to the Pauli-Lubański spin vector

$$S_a = 1/2\ e_{abcd}p^bM^{cd},$$

according to

$$S_a = sp_a,$$

can be encoded in the fact that, for some ω^A and $\pi_{A'}$,

$$p_a = \bar{\pi}_{A'}\pi_{A'}$$

and

$$M^{ab} = i\,\omega^{(A}\bar{\pi}^{B)}\epsilon^{A'B'} - i\,\epsilon^{AB}\,\bar{\omega}^{(A'}\pi^{B')} \ .$$

We regard these as field quantities, the field-point being the origin about which the angular momentum is being taken (p_a being, of course, constant). Taking $\pi_{A'}$ as likewise constant (and non-zero) we find that the required position dependence

$$M^{ab} = \overset{\circ}{M}{}^{ab} - x^a p^b + x^b p^a$$

is precisely equivalent to the position-dependence of ω^A as a solution of the twistor equation, as given above, with the same $\pi_{A'}$, i.e. to

$$\nabla_{AA'}\omega^B = -i\epsilon_A{}^B\pi_{A'}.$$

It may be remarked that the principal null directions of M^{ab} are the flagpole directions of ω^A and $\pi_{A'}$. For $\pi_{A'}$, the flagpole direction is constant, being that of the 4-momentum, while for ω^A the directions are those of a Robinson congruence - whose projections into a Euclidean 3-space of constant time are tangent to the family of circles of a stereographically projected system of Clifford parallels on S^4.

For a general physical system with 4-momentum p_a and angular momentum

$$M^{ab} = \bar{\mu}^{AB}\epsilon^{A'B'} + \epsilon^{AB}\mu^{A'B'},$$

the dependence on position (i.e. on choice of origin) is still twistorial, being defined by that of the spinor parts of the angular momentum twistor

$$A_{\alpha\beta} = \begin{pmatrix} 0 & p_A{}^{B'} \\ p^{A'}{}_B & -i\mu^{A'B'} \end{pmatrix}$$

This is subject to the Hermiticity condition

$$A_{\alpha\beta}I^{\beta\gamma} = \bar{A}^{\gamma\beta}I_{\beta\alpha}$$

where the infinity twistors $I^{\alpha\beta}$, $I_{\alpha\beta}$ are defined by

$$I^{\alpha\beta} = \begin{pmatrix} \epsilon^{AB} & 0 \\ 0 & 0 \end{pmatrix}, \qquad I_{\alpha\beta} = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon^{A'B'} \end{pmatrix}$$

and where the rules for forming the complex conjugate of a higher-valence twistor follow from that for a twistor of valence one, whereby complex conjugation sends

$$Z^\alpha = (\omega^A, \pi_{A'})$$

 Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

to the dual twistor

$$\bar{Z}_\alpha = (\bar{\pi}_A, \bar{\omega}^{A'}).$$

The corresponding α-plane is sent to its complex conjugate β-plane. We note that the rest-mass m satisfies

$$m^2 = -1/2\ A_{\alpha\beta}\bar{A}^{\alpha\beta}.$$

The Hermitian form $Z^\alpha\bar{Z}_\alpha$ has signature (++--) and in fact has the physical interpretation

$$Z^\alpha\bar{Z}_\alpha = 2s,$$

where the helicity s is the proportionality factor in $S_a = sp_a$. When $s = 0$, Z^α is a null twistor and, as we have seen, the α-plane determined by Z^α contains a real null straight line: the world-line of the massless particle whose momentum and angular momentum are p_a and M^{ab}. When $s \neq 0$ no such unique world-line exists. When $\pi_{A'} = 0$ (and $\omega^A \neq 0$) we get a limiting world-line which is a generator of $\mathscr{I}$.

5. Twistor Concepts for Curved Space-Time

It has been remarked that the twistor equation is conformally invariant, i.e. under conformal rescaling

$$g_{ab} \mapsto \hat{g}_{ab} = \Omega^2 g_{ab}$$

solutions ω^A of the twistor equation remain solutions when the Christoffel-Levi Civita connection ∇_a is replaced by the corresponding $\hat{\nabla}_a$. The corresponding $\pi_{A'}$ is modified, however, by the addition of a term involving ω^A, and we have

$$\hat{\omega}^A = \omega^A, \quad \hat{\pi}_{A'} = \pi_{A'} + i\,\omega^A\Omega^{-1}\nabla_{AA'}\Omega.$$

This shows us how to extend the flat-space twistor concept to any conformally flat curved space-time - at least in any simply-connected region.

For a general curved space-time $\mathscr{M}$, the situation is much less clear. Several different concepts of twistor have arisen, e.g. the local twistor, 2-surface twistor, 3-surface twistor, hypersurface twistor and asymptotic twistor concepts. None of these is universally satisfactory, but each is of value within the context in which it applies. In order to systematize this somewhat bewildering multiplicity, I have classified these various concepts in Table 1. In each case the concept arises as an adaptation of either the "α-plane definition" or else the "twistor equation definition" of a twistor, the two columns distinguishing these two different approaches.

Table 1

Submanifold referred to	α-plane definition	twistor equation definition
point	—	local twistor: $(\omega^A, \pi_{A'})$ at point $\hat{\omega}^A = \omega^A$, $\hat{\pi}_{A'} = \pi_{A'} + i\,\omega^A \Omega^{-1} \nabla_{AA'} \Omega$
(open) curve (tangent t^a)	—	local twistor transport: $t^a \nabla_a \omega^B = -i\, t^{BA'} \pi_{A'}$; $t^a \nabla_a \pi_{B'} = -i\, t^a P_{aBB'} \omega^B$; $(P_{ab} = \frac{1}{12} R\, g_{ab} - \frac{1}{2} R_{ab})$
2-surface $\mathcal{S}$ (spacelike, $\cong S^2$)	Spinor $\pi_{A'}$ at a point of $\mathbb{C}\mathcal{S}$	2-surface twistor: solution of $\eth' \omega^0 = \sigma' \omega^1$, $\eth \omega^1 = \sigma \omega^0$
hypersurface $\mathcal{H}$ (normal n^a)	hypersurface twistor: α-curve, with $\pi_{A'}$-scaling, on $\mathbb{C}\mathcal{H}$ $\pi_{A'} n^{AA'} \pi_{B'} \nabla_A{}^{B'} \pi_{C'} = 0$ asymptotic twistor: special case when $\mathcal{H} = \mathcal{I}$.	3-surface twistor: solution of $n^{A'}(A \nabla^B{}_{A'} \omega^C) = 0$ (needs magnetic part of Weyl tensor and its normal derivative both to vanish at $\mathcal{H}$)
whole space-time (complex)	"non-linear graviton construction" requires $\tilde{\Psi}_{A'B'C'D'} = 0$ $(C_{abcd} = \Psi_{ABCD} \varepsilon_{A'B'} \varepsilon_{C'D'} + \varepsilon_{AB} \varepsilon_{CD} \tilde{\Psi}_{A'B'C'D'})$	googly??

 Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

The various rows of the table refer to the different dimensions of submanifolds with respect to which the twistor definition refers. With respect to a point, the twistor equation definition provides the concept of a local twistor $Z^\alpha = (\omega^A, \pi_{A'})$ which is replaced, under conformal rescaling of the space-time metric, by $Z^\alpha = (\hat{\omega}^A, \hat{\pi}_{A'})$ as given by the above expressions. With respect to an (open) curve, the twistor equation reduces to local twistor transport of local twistors along the curve, which is a conformally invariant propagation. The definition is such as to agree with the behaviour of the ω^A and $\pi_{A'}$ fields in the case when $\mathcal{M}$ is conformally flat. Without conformal flatness, local twistor transport is not integrable, so the twistor concept does not globalize throughout $\mathcal{M}$. The local twistor curvature involves both the Weyl tensor and its (contracted) covariant derivative. Local twistors can be used to provide a Yang-Mills theory for conformal gravity, the Yang-Mills current being the Bach tensor (see Merkulov 1984). They also lead to a systematic procedure for generating conformally invariant objects and operations in four-dimensional space-time manifolds.

Neither with respect to a point nor a curve do we get an α-plane concept of a twistor, since the dimensions are such that an α-plane would, in general, not intersect the given submanifold. However, with respect to a submanifold $\mathcal{S}$ of dimension two, the α-plane concept of a twistor becomes relevant. For an analytic $\mathcal{S}$ in $\mathbb{M}$, its complexification $\mathbb{C}\mathcal{S}$ will (generally) meet some open family of α-planes lying in $\mathbb{CM}$. (The symbol "$\mathbb{C}$" used in this way always means "complexified". Here $\mathbb{C}\mathcal{S}$ would be some complex "thickening" of $\mathcal{S}$ in $\mathbb{CM}$, i.e. a complex two-dimensional submanifold of $\mathbb{CM}$ whose intersection with $\mathbb{M}$ is $\mathcal{S}$.) At any point of intersection of $\mathbb{C}\mathcal{S}$ with an α-plane, the spinor $\pi_{A'}$ would determine the α-plane, so we can think of a twistor (belonging to the given open family) as being defined by a point of $\mathbb{C}\mathcal{S}$ together with a spinor $\pi_{A'}$ at that point. This definition carries over to an (analytic) surface $\mathcal{S}$ in a general (analytic) space-time $\mathcal{M}$. (Some subtleties arise when the flagpole direction of $\pi_{A'}$ becomes close to being either normal or tangent to $\mathbb{C}\mathcal{S}$, but I shall not go into these here). For a spacelike 2-surface $\mathcal{S}$ (with S^2 topology) the twistor-equation definition will also apply, and it leads us to the definition of a 2-surface twistor. This important case, and its significance for general relativity, will be discussed in the next two sections.

Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

For a three-dimensional analytic hypersurface $\mathcal{H}$ we have a concept of twistor arising both from the α-plane definition and from the twistor-equation definition, providing hypersurface twistors and Tod's (1983,1984) 3-surface twistors, respectively. In complexified (conformally) flat space-time, α-planes exist and these intersect $\mathbb{C}\mathcal{H}$ in complex curves called α-curves defined by the equation

$$\pi^{A'} n^{A'B'} \pi_{B'} \nabla_{AA'} \pi_{C'} = 0,$$

where n^a is normal to $\mathbb{C}\mathcal{H}$ and $\pi^{A'} n^{A'B'} \pi_{B'}$ is tangent to the curve. The α-curve defines the α-plane, and the twistor scaling is fixed by $\pi_{A'}$. In general there are no α-planes in $\mathbb{C}\mathcal{M}$, but we can still use α-curves (with $\pi_{A'}$-scaling) to provide a twistor concept - in fact, the hypersurface twistor concept (Penrose 1975, Penrose and Rindler 1986). The space $\mathcal{T}^\bullet(\mathcal{H})$ of hypersurface twistors for $\mathcal{H}$ is a complex manifold, but it does not generally have a linear structure. This is a general feature of the α-plane definition. The twistor-equation definition, on the other hand, always provides a linear structure, but only applies to certain special 3-surfaces, namely those in which the magnetic part of the Weyl tensor vanishes and the normal derivative of this magnetic part also vanishes (this latter being equivalent, in the case of a vacuum space-time, to the condition that the curl of the electric part vanishes). These are the integrability conditions for the 3-surface twistor equation

$$n^{A'(A} \nabla^{B}_{A'} \omega^{C)} = 0,$$

this equation being the restriction of the twistor equation to directions lying within $\mathcal{H}$ (see Tod 1983). When a complete 4-parameter family of solutions (3-surface twistors) exists, we can identify the hypersurface twistors in terms of the zero-loci of 3-surface twistors. More directly, we can see that the complex conjugates of hypersurface twistors are naturally dual to solutions of the 3-surface twistor equation, and this leads to a linear structure to the space of hypersurface twistors in this case.

Despite the lack of linear structure, in the general case, hypersurface twistors are still of value. As the hypersurface $\mathcal{H}$ is moved within $\mathcal{M}$, the structure of $\mathcal{T}^\bullet(\mathcal{H})$ evolves. This can be related to the initial value problem and Hamiltonian evolution for Einstein's equations (Mason 1986 cf. also LeBrun 1984). The definition of hypersurface twistors is conformally invariant, and so it can be applied

Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

to $\mathscr{I}$ in the case when $\mathscr{M}$ is (appropriately) asymptotically flat. The twistors so obtained are called asymptotic twistors (Penrose and MacCallum 1972). In this case the α-curves become null geodesics on $\mathbb{C}\mathscr{I}$.

When applied to the whole space-time $\mathscr{M}$, neither the α-plane definition nor the twistor-equation definition directly applies in general. Defining an α-plane (or "α-surface") in $\mathbb{C}\mathscr{M}$ to be a complex 2-surface whose tangent vectors at any one point all have the form $\lambda^A\pi^{A'}$ for fixed $\pi^{A'}$ and varying λ^A, we find by a Lie bracket argument, that

$$\Psi_{A'B'C'D'}\pi^{A'}\pi^{B'}\pi^{C'}\pi^{D'} = 0.$$

Thus, if there is an α-plane corresponding to each $\pi_{A'}$ through each point of $\mathbb{C}\mathscr{M}$ (giving a complex three-parameter family of α-planes in all) then we must have

$$\Psi_{A'B'C'D'} = 0$$

throughout $\mathbb{C}\mathscr{M}$, which is to say that $\mathbb{C}\mathscr{M}$ is right-conformally flat. (The complex Weyl curvature is given by $C_{abcd} = \Psi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\bar{\Psi}_{A'B'C'D'}$ with Ψ_{ABCD} and $\bar{\Psi}_{A'B'C'D'}$ both symmetric, and describing the anti-self-dual and self-dual parts of the conformal curvature, respectively.) For $\mathbb{C}\mathscr{M}$ to be actually the complexification of a Lorentzian space-time $\mathscr{M}$, this would require $\mathscr{M}$ to be conformally flat. I shall abuse the notation, somewhat, here and simply use $\mathbb{C}\mathscr{M}$ to denote a complex space-time.

For any right-conformally flat $\mathbb{C}\mathscr{M}$, the projective twistor space $\mathbb{P}\mathscr{T}$, whose points represent α-planes in $\mathbb{C}\mathscr{M}$ (considered only locally in $\mathbb{C}\mathscr{M}$), and the non-projective twistor space $\mathscr{T}$, for which the scaling of the $\pi_{A'}$-spinors is also included, are well defined and can be used to reconstruct the (complex) conformal metric of $\mathbb{C}\mathscr{M}$. When $\mathbb{C}\mathscr{M}$ is also Ricci-flat - so $\mathbb{C}\mathscr{M}$ is right flat, i.e. a general anti-self-dual solution of the Einstein fluid equations - the metric itself emerges from the structure of $\mathscr{T}$, which is now strengthened by the existence of a holomorphic fibration

$$\mathscr{T} \to \mathbb{C}^2 - (0),$$

(where $\mathbb{C}^2$ stands for the space of $\pi_{A'}$) whose fibres are complex symplectic 2-manifolds (i.e. with a complex area form) with appropriate homogeneity properties. The construction of $\mathscr{T}$ requires the solution of no algebraic or differential equations whatever, but the construction of the points of $\mathbb{C}\mathscr{M}$ from $\mathscr{T}$ requires finding the family of holomorphic cross-sections of the above fibration, each cross-section giving a point

of $\mathbb{C}\mathcal{M}$, which may be hard to do explicitly. This is the "non-linear graviton" (or "leg-break") construction (Penrose 1976; see also Ward 1978, Tod and Ward 1979, Hitchin 1979, 1982, for explicit examples, and Ward 1980 for the inclusion of a cosmological constant).

There are now some good indications that a corresponding construction for the self-dual Einstein equations can be found in terms of a twistor space (as opposed to a dual twistor space - which could be trivially obtained by the previous construction merely by a change of conventions). Such a construction is referred to as a "googly" (in view of a cricketing analogy). The googly construction appears to have features in common with a twistor-equation definition rather than an α-plane definition of a twistor, but the matter is not yet wholely clear. The idea is that both the googly and leg-break structures of $\mathcal{T}$ can be imposed simultaneously, so that there is then the possibility of constructing general solutions of the full (i.e. neither anti-self-dual nor self-dual) Einstein equations. All this is work in progress.

6. Quasi-Local Mass and Angular Momentum

Let us return to the case of a 2-surface $\mathcal{S}$, which we take to be spacelike, with S^2 topology and we consider the twistor-equation definition of a twistor. Choosing a spin-frame o^A, ι^A whose flagpole directions are normal to $\mathcal{S}$, the required tangential part of the twistor equation becomes

$$\eth' \omega^0 = \sigma' \omega^1, \quad \eth \omega^1 = \sigma \omega^0.$$

It turns out (by an argument involving the Atiyah-Singer index theorem) that there are always four independent solutions to this pair of equations and, in general, no more than four. (Jeffryes has constructed examples where there are more than four, but it is not known whether this can ever happen in physically realistic situations cf. also Jeffryes 1984.) When four-dimensional, the solution space of these equations is referred to as the 2-surface twistor space $\mathbb{T}^\alpha(\mathcal{S})$, and a solution ω^A is referred to as a 2-surface twistor on $\mathcal{S}$, denoted $Z^\alpha \in \mathbb{T}^\alpha(\mathcal{S})$.

The main significance of 2-surface twistors is that they can be used as spin-lowering operators on $\mathcal{S}$ which lower the "spin 2" curvature tensor, in two steps, to a "spin 1" Maxwell-type field on $\mathcal{S}$. This leads to a definition of the total mass-momentum and angular momentum within $\mathcal{S}$

Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

in terms of "charge integrals" for the Maxwell-type field (Penrose 1982, 1984, Penrose and Rindler 1986). This complex of quantities constitutes the 2-surface angular momentum twistor $A_{\alpha\beta} \in \mathbb{T}_{(\alpha\beta)}(\mathcal{S})$ given by

$$A_{\alpha\beta}{}^{Z}_{1}{}^{\alpha}{}^{Z}_{2}{}^{\beta} = \frac{-1}{16\pi} \oint \eta \, R_{abCC'DD'} \, {}^{C}_{1}\omega \, {}^{D}_{2}\omega \, \epsilon^{C'D'} dx^{a} \wedge dx^{b}$$

integrated over $\mathcal{S}$, where η varies over $\mathcal{S}$ in direct proportion to the determinant

$$\begin{vmatrix} \omega_1^0 & \omega_2^0 & \omega_3^0 & \omega_4^0 \\ \omega_1^1 & \omega_2^1 & \omega_3^1 & \omega_4^1 \\ \eth\omega_1^0 & \eth\omega_2^0 & \eth\omega_3^0 & \eth\omega_4^0 \\ \eth'\omega_1^1 & \eth'\omega_2^1 & \eth'\omega_3^1 & \eth'\omega_4^1 \end{vmatrix}$$

The numbers written beneath the symbols Z and ω refer to (corresponding) different arbitrarily chosen 2-surface twistors. The η-factor can be taken to be unity in many circumstances (i.e. wherever the determinant is constant over $\mathcal{S}$), but in the general case there is an uncertainty as to the appropriate way in which η should be normalized (corresponding to the most appropriate choice of 2-surface Levi-Civita twistor $\epsilon_{\alpha\beta\gamma\delta}$). There is an uncertainty, also, as to the most appropriate definition for a complex conjugation map

$$Z^{\alpha} \mapsto \overline{Z}_{\alpha}$$

and for an infinity twistor $I^{\alpha\beta} = -I^{\beta\alpha}$. (Such an infinity twistor would have to be of a "de Sitter" rather than a "Poincaré" type, in general, i.e. it would have to have matrix rank <u>four</u> rather than <u>two</u>.) Modulo such uncertainties we can define the total <u>rest-mass</u> within $\mathscr{S}$ using the formula $m^2 = -1/2\ A_{\alpha\beta}\bar{A}^{\alpha\beta}$ as given above (or else using $m^4 = 1/6\ A_{\alpha\beta}A_{\gamma\delta}A_{\lambda\mu}A_{\rho\sigma}\epsilon^{\alpha\gamma\lambda\rho}\epsilon^{\beta\delta\mu\sigma}$, an alternative suggested by Tod).

For an asymptotically flat space-time, the quasi-local 4-momentum reduces, at space-like infinity, to that of Arnowitt, Deser and Misner, and the quasi-local angular momentum reduces to that of Ashtekar and Hansen (Shaw 1983, 1984). At null infinity the quasi-local 4-momentum reduces to that of Bondi and Scahs, whereas the quasi-local angular momentum provides a new and rather more satisfactory definition than had been obtainable previously (Penrose 1982, Penrose and Rindler 1986). it has still not proved possible to find a completely satisfactory flux law for angular momentum at null infinity, however (Shaw 1984).

7. <u>Examples of Quasi-Local Mass</u>

The quasi-local mass has been calculated for a number of particular examples, mainly by Tod (1983, 1985) who has employed the device of embedding $\mathscr{S}$ in a conformally flat space-time in such a way that the induced metric on $\mathscr{S}$ remains unchanged and so do the extrinsic curvature quantities at $\mathscr{S}$. (These extrinsic quantities are simply the spin-coefficients $\sigma,\sigma',\rho,\rho'$ at $\mathscr{S}$.) A surface $\mathscr{S}$ for which it is possible to do this is called <u>uncontorted</u>; otherwise $\mathscr{S}$ is called <u>contorted</u>. Uncontorted surfaces have the advantage that we can choose

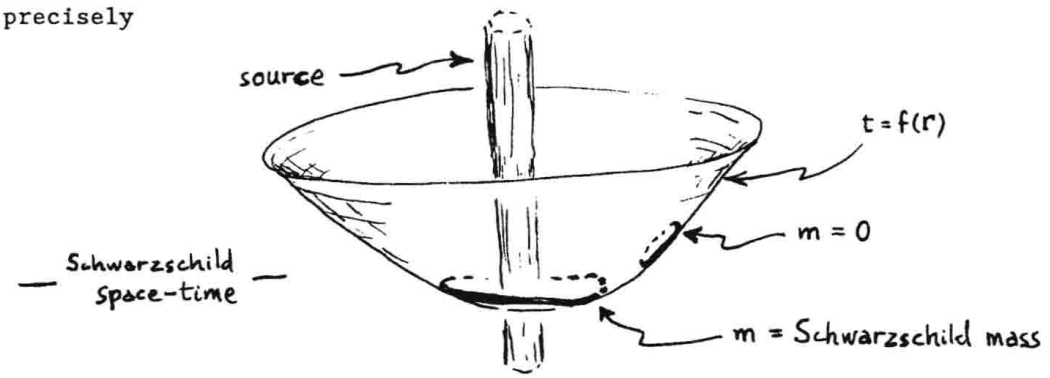
Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

$\eta = 1$, since the relevant determinant is always constant over $\mathcal{S}$. Moreover, the complex conjugate map $Z^\alpha \mapsto \overline{Z}_\alpha$ is naturally defined in these cases, so that m^2 aquires an unambiguous definition. Calculations are facilitated by the fact that the 2-surface twistors are all induced from solutions of the twistor equation in the conformally flat embedding space.

In the Schwarzschild space-time, any 2-surface $\mathcal{S}$ lying in a hypersurface of revolution

$$t = f(r)$$

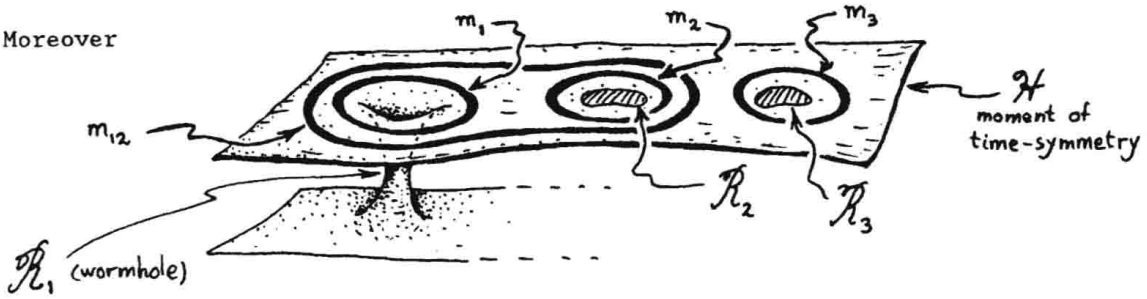
is uncontorted (this hypersurface admitting 3-surface twistors). It turns out $m = 0$ if such an $\mathcal{S}$ does not link the source, while m becomes precisely



the Schwarzschild mass whenever $\mathcal{S}$ does link the source (once). For the Reissner-Nordstrom space-time a similar result holds except that there is a correction to the mass (as expected) arising from the electromagnetic energy density (Tod 1983).

Another of Tod's examples arises when $\mathcal{S}$ lies in a spacelike hypersurface $\mathcal{H}$ which is a spatially conformally flat moment of time-symmetry. Suppose that on $\mathcal{H}$, there are localized regions of matter $\mathcal{R}_1$, $\mathcal{R}_2$, ... (or else wormholes connecting $\mathcal{H}$ to a second copy of itself). Again $\mathcal{S}$ is uncontorted (and $\mathcal{H}$ admits 3-surface twistors) and,

for $\mathcal{S}$ entirely in the vacuum region, m turns out to be a function only of the topological linking that $\mathcal{S}$ has with the various source regions. Moreover



if m_1, m_2 and m_{12} are the respective masses obtained when the regions $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_1 \smile \mathcal{R}_2$ are surrounded, then the physically reasonable inequality

$$m_{12} < m_1 + m_2$$

is obtained. In the limit of large separations the difference

$$m_{12} - m_1 - m_2$$

agrees with the expected negative contribution from the Newtonian potential energy. This is very gratifying, and it is further reassuring that (as Jeffryes 1986 has recently shown) in the Newtonian limit of a general system, the quasi-local mass expression correctly incorporates the mass due to Newtonian potential energy.

Tod has also shown that for a static "black hole in a cage" - where the cage is to distort the black hole out of spherical shape - with $\mathcal{S}$ chosen at the hole's horizon, of area A, there is an inequality $m \leqslant (A/16\pi)^{1/2}$, equality holding only in the spherically symmetrical case. This also seems physically reasonable, there being an energy due to the distortion of the hole out of spherical shape.

630 Twistors in general relativity. *General relativity and gravitation. (Proc. 11th Int. Conf. General Relativity and Gravitation, 1986).* Ed. M. A. H. MacCallum, 158–76. 1987

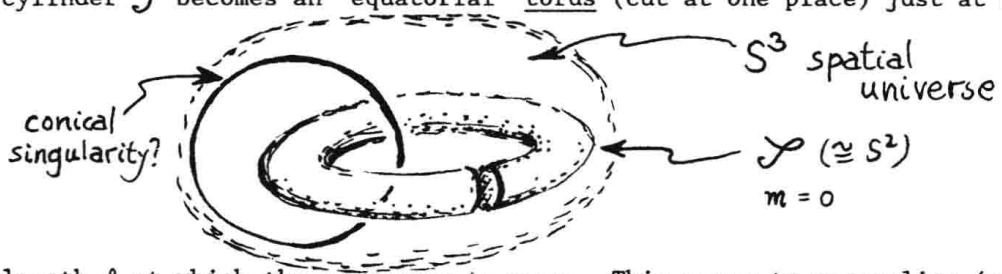
Recently, Tod (1986) (and also Jeffryes and Reula) have examined cylindrically symmetrical space-times, where $\mathscr{I}$ is a cylinder with flat ends. A dependence of the mass on the length ℓ of the cylinder is found which falls into three different categories:

Hyperbolic: exponential dependence on ℓ

Elliptic: sinusoidal dependence on ℓ

Parabolic: linear dependence on ℓ.

At first, such dependence seems puzzling, particularly in the elliptic case, since the mass would seem to go negative after a critical length is reached. However, these results may be compared with earlier calculations by Tod concerning the Friedman-Robertson-Walker models (Tod 1983). There a similar behaviour was found and taken to be "physically reasonable". The puzzling elliptic case now occurs with k = +1 models, but here the cylinder $\mathscr{I}$ becomes an "equatorial" torus (cut at one place) just at the



length ℓ at which the mass goes to zero. This seems to generalize (and be compatible with) the result that the total mass of a closed universe should be zero.

One might envisage a similar situation but where the mass is allowed to become negative, by introducing a conical singularity linking the torus, so that the mass goes negative before the torus closes. However, this introduces negative mass outside $\mathscr{I}$, which must be regarded as physically unreasonable. There is, however, a moral to this

which is a warning to anyone attempting to prove a positivity theorem for quasi-local amss: the space must be physically reasonable both inside and outside of $\mathcal{S}$.

Finally, the calculations of Kelly, Tod and Woodhouse (1986) concerning m for small spheres and ellipsoids may be mentioned. If $\mathcal{S}$ is such a small ellipsoid, its linear size being characterized by a parameter r, they find that at order r^{-3} there is the expected contribution from the energy-momentum tensor times volume. In vacuum the first possibility of a non-zero contribution occurs at order r^{-5}. Without the η-factor in the quasi-local mass definition anomalies arise ($m^2 < 0$) when $\mathcal{S}$ is contorted. But these miraculously vanish when η is included. Further work is in progress.

REFERENCES

Aspect, A., Grangier, P. & Roger, G. (1981). Phys. Rev. Lett. 47, 460

Freedman, S.J. & Clauser, J.F. (1972). Phys. Rev. Lett. 28, 938

Hitchin, N.J. (1979). Polygons and gravitons, *Math. Proc. Camb. Phil. Soc.* 85, 465-76

Hitchin, N.J. (1982). Complex manifolds and Einstein's equations, in *Twistor Geometry and Non-Linear Systems* ed. H.D. Doebner and T.D. Palev (Springer-Verlag, Berlin), pp. 73-99

Huggett, S. & Tod, K.P. (1985). Introduction to Twistor Theory, London Math. Soc. Lecture Note Series (Cambridge University Press, Cambridge)

Jeffryes, B.P. (1984b). Two-surface twistors and conformal embedding in *Asymptotic Behaviour of Mass and Spacetime Geometry* , Proc. Corvalis, Oregon 1983, ed. F.J. Flaherty (Springer-Verlag, Berlin)

Kelly, R., Tod, K.P. & Woodhouse, N.M.J. (1986). Classical and Quantum Gravity (to appear)

LeBrun, C.R. (1984). Twistor CR-manifolds and three-dimensional conformal geometry, *Trans. Amer. Math. Soc. 278, 209-31*

Mason, L. (1986). Twistors in curved space time, D.Phil. thesis, Oxford

Merkulov, S.A. (1984). A conformally invariant theory of gravitation and electromagnetism, Class. Quantum Grav. , 1, 349-54

Penrose, R. (1975). Twistor theory: its aims and achievements, in *Quantum Gravity, an Oxford Symposium* , eds. C.J. Isham, R. Penrose and D.W. Sciama (Oxford University Press, Oxford)

Penrose, R. (1976). Nonlinear gravitons and curved twistor theory, *Gen. Rel. Grav.* 7,, 31-52; The non-linear graviton, ibid., 171-6

Penrose, R. (1982). Quasi-local mass and angular momentum in general relativity, *Proc. Roy. Soc. London* A381, 53-63

Penrose, R. (1984). Integrals for general-relativistic sources: a development from Maxwell's electromagnetic theory, in *Maxwell Symposium Volume* , ed. M.S. Berger (North-Holland, Amsterdam)

Penrose, R. & MacCallum, M.A.H. (1972). Twistor theory: an approach to to the quantization of fields and space-time, *Phys. Repts.* 6C, 241-315

Penrose, R. & Rindler, W. (1984). Spinors and Space-Time, Vol. 1, Two-Spinor Calculus and Relativistic Fields (Cambridge University Press, Cambridge)

Penrose, R. & Rindler, W. (1986). Spinors and Space-Time Vol. 2, Spinor and Twistor Methods in Space-Time Geometry (Cambridge University Press, Cambridge)

Shaw, W.T. (1983). Twistor theory and the energy-momentum and angular momentum of the gravitational field at spatial infinity, *Proc. Roy. Soc. London* A390, 191-215

Shaw, W.T. (1984). Twistors, asymptotic symmetries and conservation laws at null and spatial infinity, in *Asymptotic Behaviour of Mass and Spacetime Geometry* , Proc. Corvalis, Oregon 1983, ed. F.J. Flaherty (Springer-Verlag, Berlin)

Tod, K.P. (1983). some examples of Penrose's quasi-local mass construction, *Proc. Roy. Soc. London* A388, 457-77

Tod, K.P. (1984). Three-surface twistors and conformal embedding, *Gen. Rel. Grav.* 16, 435-43

Tod, K.P. and Ward, R.S. (1979). Self-dual metrics with self-dual Killing vectors, *Proc. Roy. Soc. London* A368, 411-27

Ward, R.S. (1978). A class of self-dual solutions of Einstein's equations, *Proc. Roy. Soc.* A363, 289-95

Ward, R.S. (1980). Self-dual space-times with cosmological constant, *Comm. Math. Phys.* 78, 1-17

174
NEWTON, QUANTUM THEORY AND REALITY

Three hundred years of gravitation. Eds S. W. Hawking and W. Israel, 17–49. Cambridge University Press, 1987

This is one of my few excursions into the history/philosophy of science. At the proof stage, Nick Woodhouse brought to my attention the manuscript fragment referred to on p. [49], and I was struck by the fact that this appears to give support to my contention that Newton must indeed have been, at heart, a 'relativist'. On the other hand, I now understand that Hooke had observed, before Newton, what we now refer to as 'Newton's rings', in partial contradiction with my comments towards the end of p. [17].

 Newton, quantum theory and reality. *Three hundred years of gravitation.* Eds S. W. Hawking and W. Israel, 17–49. 1987

3

Newton, quantum theory and reality

ROGER PENROSE

3.1 Newton's corpuscular–undulatory view of light

The supreme stature of Newton as a scientist cannot be doubted. As an experimental physicist, he had superb natural skill, and was profoundly ingenious, as well as being exceptionally careful in the construction and execution of his experiments. As a mathematician he possessed extraordinary power, and, indeed, had few mathematical peers over the whole of history. In fact, it was very necessary for him, in developing his scientific theories, to explore and innovate in mathematical ways, and he uncovered many deep mathematical truths – the fundamental theorem of calculus being but one. The *Principia* itself, whose three-hundredth anniversary I am most honoured to be able to celebrate with this essay, contained among its many propositions, multitudes of mathematical gems and numerous profound mathematical ideas. In addition to these scientific qualities, and most importantly for the development of his science, Newton also possessed superlative instincts into the workings of Nature – instincts which often fell hardly short of miraculous.

In one particular instance, concerning Newton's preference for a corpuscular theory of light to a wave theory, the extent of his insight is most intriguing – and perhaps somewhat debatable. The experimental evidence in Newton's day was inconclusive, yet it would seem that the balance of that evidence should have been in support of a wave theory of light rather than a corpuscular one. Ironically, it was Newton's own discovery and painstaking analysis of an optical interference phenomenon – Newton's rings – which should have provided the strongest evidence in favour of the opposing wave theory! (Other types of interference had been observed earlier by Grimaldi.) Yet Newton stuck doggedly to a corpuscular view, and was led to a picture of light quanta possessing both particle-like and wave-like properties. Can

this apparent anticipation of modern quantum theory – a theory which was not formulated until over two hundred years later – be other than the merest fluke, resulting perhaps from Newton's obstinate refusal to give in to his adversaries and accept their wave picture of light?

In his preface to the 1952 Dover edition to Newton's *Opticks*, I. Bernard Cohen rightly warns us not to make too much of Newton's apparent anticipation of our present quantum-mechanical concepts. Yet, one cannot but be struck by the fact that this master among scientists found himself driven to a view of the physical nature of light that was not only seemingly almost absurd, with its curious intermixture of corpuscular and undulatory properties, but absurd in ways which seem close to the absurdities of modern quantum theory! Is not this an indication that Newton possessed insights of such power that they even gave him hints of revolutionary developments not to occur until over two centuries later? On the other hand, perhaps the similarity with quantum mechanics is indeed illusory. Is there really much detailed correspondence between Newton's undulatory 'fits of easy transmission and reflection' when his corpuscles of light interact with the material of a refractive medium, and the wave-like aspects of photons in their interactions with matter, as we understand them today? I would accept that these analogies should not be pressed too far, and that the correspondence between Newton's picture of light and that presented by modern quantum mechanics would not stand up to much detailed analysis. I can believe, also, that Newton possessed the mathematical wherewithal, so that in principle he could have pressed forward with an entirely wave picture of light, had he wished to do so, which would have been well in accord with the observations of his day. Yet I think that it is likely that Newton did indeed have powerful reasons for rejecting a purely wave picture of light; and there was evidence available even at that time, vaguely discernible to a scientist of Newton's sensitivity, for a view of reality in which undulatory and corpuscular ingredients of reality co-exist at a fundamental level. However, such reasons would have to have been ones that Newton could not easily have articulated – or perhaps would not have wished to do so, owing to his avowed dislike of purely 'hypothetical' contentions.

3.2 Stated reasons for rejection of wave theory

Let us first briefly examine Newton's stated reasons for rejecting the wave theory. The main one seems to have been the rectilinear propagation of light, and the fact that light does not bend into the shadow of an obscuring body. As we now know, from the phenomenon of destructive interference

discovered by Thomas Young a century later, an effective rectilinear propagation does indeed occur for wave propagation when the wavelength is short, such as that which occurs with visible light – and light does indeed bend slightly into the shadow. We do not blame Newton and his contemporaries for not knowing this. Yet I suspect that if Newton had preferred the wave theory for other reasons, it might not have been beyond him to suggest something of the nature of destructive interference. (The wavelengths of visible light were, in effect, known to him from his 'Newton's rings' observations.) However, his preferences were for a corpuscular theory, which rectilinear propagation seemed to support, so he did not find himself compelled to examine this matter more deeply.

Newton's other stated reason for rejecting the wave theory of light was that light has 'sides' – which was Newton's term for linear polarization. The observations of Erasmus Bartholin and Christiaan Huygens concerning the phenomenon of double refraction in calcite, or Iceland spar, had led Newton to conclude that a beam of light consists of components which are not rotationally symmetric (i.e., they are polarized). This is, indeed, quite incompatible with the type of longitudinal compression wave that Huygens had suggested for light. However, a transverse oscillation, as was later suggested by Young, would fit the bill. Apparently it occurred to neither Huygens nor Newton that light might be a transverse oscillation. One can again speculate that if Newton had had other strong reasons for preferring a wave theory of light, he might perhaps have been driven to think of the idea of transverse oscillations as the explanation of 'sides'. However, he had not and he was not.

These two arguments for rejecting the wave theory of light, though undoubtedly impressive at the time, must be counted as 'wrong'; so if this was actually all that Newton had to go on, we should, indeed, have to consider that it was a 'fluke' that he found himself driven to a combined corpuscular–undulatory theory of light. However, it seems to me to be likely that Newton may well have had various other reasons for disliking an entirely undulatory theory of light. One can only surmise what some of these might have been. One possibility that has often been suggested is that Newton favoured an atomistic view of Nature. Perhaps there is some force to such an argument. If we are to believe that light is a wave motion, then surely it must be a wave in some medium. On the atomistic view, this medium must itself be composed of particles. This gives light a character secondary to that of the particles of the medium. Yet, there is perhaps something primitive about light that argues against it having such a

secondary nature. If new particles are to be invoked in order to compose such a hypothetical all-pervasive medium, is it not much more economical to use such new particles as the constituents merely of the light itself?

Newton seems to have had some difficulties in coming to terms with such an all-pervasive medium, and he discussed, at some length, the strange hypothetical properties that such a medium would need to have (in *Queries* 18–22 of *Book Three* of his *Opticks*). Even when contemplating such a medium, he preferred to have light constituted separately as particles of its own. (Influences in the medium were imagined somehow to be able to overtake the light corpuscles.) He also had some idea that an explanation of motion under *gravity* might be found in terms of 'refraction' by such an all-pervasive medium, and he presented a picture which even has something in common with the ideas of general relativity! (Perhaps, instead, Newton's ideas were foreshadowing De Broglie's realization that matter itself possessed wave-like qualities. Take your pick!) Newton appeared to have no qualms about suggesting such a view of the nature of gravity, while at the same time holding to his picture of an inverse square law of force holding between all pairs of particles in the universe. No doubt he regarded all these pictures as highly provisional. He admitted to not knowing what the nature of such a medium might be (*Query* 21). The medium would need to have some strange and powerful properties; and I suspect that Newton's view was that the medium, if it existed, could not be material in any ordinary sense, but would have to be of some mysterious substance quite different from that of matter. Perhaps this was the 'absolute space' that Newton was really groping for. Perhaps, indeed, if the modern ideas could have been explained to him, Newton might not have been averse to viewing this medium in terms of Einstein's picture of a 'flexible' spacetime continuum!

None of this implies that light must be corpuscular. Ordinary atomism might require merely that the medium, if it exists, be corpuscular. It is just economy, or perhaps aesthetic requirements, which, together with atomism, would imply that light must be corpuscular. I wish to suggest that there may, instead, have been another quite different motivation underlying Newton's preference for a corpuscular view of light. This is the motivation from Galilean relativity.

3.3 Newton and relativity

The reader may well object that there is not much evidence that Newton was at heart a dynamical 'relativist'. He was certainly well aware that his dynamical laws are invariant under change to a uniformly moving frame (as

 Newton, quantum theory and reality. *Three hundred years of gravitation.* Eds S. W. Hawking and W. Israel, 17–49. 1987

well as under shift of origin and under rotation). This much he made clear in his *Principia*. But what evidence is there that Newton at any time shared Einstein's *conviction* that physics *must* be invariant under change to uniform motion? The common view seems to be that, on the contrary, 'Newton believed in absolute space' – with its decidedly non-relativistic connotations. If space (as opposed to spacetime) is indeed absolute, so that it is meaningful to talk about the same point in space at two different times, then there is an absolutely preferred state of rest. It then becomes a 'fluke' that Newton's dynamical laws are Galilei-invariant. There is no need to believe that this fluke should continue to hold when additional ingredients of physics are added to the basic dynamical laws. One such additional ingredient is the velocity of the medium in which light travels, if it indeed be the case that light merely consists of wave motion in some all-pervading medium. If we insist that light is merely a wave motion, then we must give up Galilean relativity.

Of course, Lorentz, Einstein and Poincaré have now taught us that there is a way out of this. Light can still consist of a wave motion – this time, the waves of oscillating electric and magnetic fields of the Maxwell theory – and invariance under change to a uniformly moving frame can be maintained. However, this requires a drastic revision of dynamical laws (for speeds approaching that of light) and also an overturning of one's very notions of space and time. It is thus, as we now know, that Galilean relativity can be given up and replaced by *another* relativity – the relativity of Einstein and Poincaré.

I am certainly not attempting to suggest that Newton could have had the absurd foresight to sense the need, or even the possibility of such a revision of basic principles. However, if we do accept that this route was *not* effectively open to him, then we see that Newton was presented with a clear choice: give up Galilean relativity or stick to a corpuscular theory of light. Now what evidence is there that Newton believed 'in his bones' that some kind of dynamical relativity must hold? I contend that, despite what appears to be a common belief to the contrary, such evidence does exist. I refer the reader to Richard Westfall's fine biography of Newton (*Never at Rest*, 1980). On p. 147 Westfall describes how, in January 1665, recorded in private notes to himself in a work referred to as his *Waste Book*, the youthful Newton derived the laws of impact for two perfectly elastic bodies. (Huygen's slightly earlier treatment of this problem had not yet been published.) Newton's procedure was to consider the two bodies as a single system, and he came to the conclusion that their common centre of gravity moves inertially whether or not the bodies impinge upon one another. Realizing that every impact

can be viewed from the special frame of reference of the common centre of gravity of the two bodies, Newton appears to have used the impact behaviour in *that* frame – in which the behaviour may be regarded as 'obvious' – and then assumed *Galilean relativity* to derive the elastic impact law in the general case. This was more than twenty years before the publication of *Principia*, but Westfall remarks, concerning the consequent uniform motion of the centre of gravity, that 'Newton never forgot this conclusion. It contained the first adumbration of his third law of motion, . . .' – and it seems to have formed one of the central guiding ideas behind Newton's dynamics, despite the fact that in *Principia* Newton chose to derive his dynamics from different axiomatic principles.

In his account, Westfall does not draw particular attention to the fact that Newton was implicitly relying on Galilean invariance to derive his impact law,† although Newton's procedure cannot be used without such an assumption. Without it, one has no right to 'pass to the special frame of reference of the common centre of gravity' and to assume that the same dynamical laws must hold in that frame as in the original frame. It seems to me to be highly likely that by the age of 23, when Newton produced this argument, some strong intuition concerning the relativity of motion was already an essential ingredient of his thinking. It is, after all, a necessary ingredient of the Copernican–Galilean mode of thought that the dynamics that we employ at the surface of the Earth should be insensitive to the Earth's uniform motion about the Sun. Newton was well aware of this, of course. (It appears that, at least by 1666, Newton had read Galileo's *Dialogues*, which appeared in translation in England in 1665.) It would have

† Towards the end of his life, Galileo had claimed a solution of the problem of the dynamics of impulses. However, no account of this appears to have survived (cf. Galilei, 1638; Dover edn, p. 293). It is intriguing to speculate whether Galileo might have been guided along similar lines of reasoning.

Although Huygens's treatment of this problem seems to have been a little different from that of Newton, he also used Galilean invariance (cf. Mach, 1972). It is striking to note that Huygens did *not* believe that this Galilean invariance extended to light, in the sense that I mean it here, since he strongly promoted the wave theory! Huygens's view seems to have been that the individual particles of the medium might be subject to a Galilei-invariant dynamics even though the medium as a whole might define a rest-frame. However, this viewpoint attaches an 'accidental' rather than a 'fundamental' role to this medium – and therefore to light itself. It also presupposes that the Galilei-invariance of the local dynamics of *macroscopic* bodies is insensitive to the presence of the medium – another 'accidental' (and, moreover, highly improbable) feature. While the example of Huygens shows that it was possible for a great scientist of the time to hold to the wave theory and also to be a professed relativist, this is not totally being a relativist in the 'deep' sense that I mean here.

been natural for him to suppose that other physical properties (such as the behaviour of light) should also be governed by the same beautiful invariance principles. Perhaps, at that time, he had not explicitly formulated, even to himself, what such 'invariance principles' might entail, but my guess would be that he knew pretty well within himself what he was about. His treatment of the problem of impulses seems to me clearly to indicate that he instinctively felt that a relativity principle must be intrinsic to Nature's ways.

At roughly the same age as he studied the impulse problem, he was also becoming profoundly interested in the nature of light. On Newton's own account, he 'had the Theory of Colours' in January 1666 (Westfall, 1980, p. 156). He may even have had some inkling (as had Galileo before him – cf. Drake, 1957, p. 278, concerning a remark made in *The Assayer*) that light has some profound role to play in governing the minute structure of matter. One can readily imagine that Newton's great intuition guided him to believe strongly in some kind of relativity *principle* encompassing both the dynamics of matter and the behaviour of light. This intuition would have guided him surely to a corpuscular rather than to a wave picture of light. That would have been in the years when his profound views on the nature of light were first being formed. These years must have set the stage for all his later thinking. When at a later time in his life he became fully confronted by the experimental facts of optical interference phenomena, he would have been reluctant to give up the intuitions that he had built up during this formative time. I suggest that in his 'bones' he *knew* that light must be corpuscular, even though in later years the good reasons for these early intuitions may well have receded from him. And unlike the stated reasons that Newton later put forward in his *Opticks* in favour of a corpuscular picture of light – which, technically, were 'incorrect' – this unstated one, that a relativity principle must hold, was, I am claiming, profoundly 'correct'!

When Newton came to write his *Principia*, some twenty years after these intuitions must have been formed, he needed to place his logical development on a clear footing. To state his basic laws in a definite mathematical way, he needed the framework of 'absolute space' so as to be able to formulate things precisely. Even with the abstract mathematical machinery of today, it is not so easy to give a totally Galilean-relativistic treatment of space, time and dynamics right from the start. It can be done, but one needs an absolute spacetime, rather than an absolute space. (See Arnol'd, 1978, for an excellent modern such treatment.) Newton did take pains to be completely clear that Galilean invariance was indeed a feature of his laws (cf. Scholium 5, following Definition VIII), but this invariance was

not a necessary principle. It might even be that the technicalities involved in writing his *Principia* guided Newton away from what I believe must have been his early relativistic intuitions. Once Galilean relativity has become a *consequence* of the basic laws, rather than a basic principle, there appears to be nothing against adding to those laws something – such as an all-pervasive medium – which does not share in that Galilean invariance. Accepting all the profound influence, power and comprehensiveness of *Principia*, and accepting all its magic in deriving from such simple axioms so much of the dynamical workings of the world, one can still contemplate the case that its very strength of mathematical structure could have obscured a youthful insight – an insight that had to wait nearly two and one half centuries to be profoundly reinstated in a new form by Einstein in 1905: that one should take relativity as a *principle*, rather than as a seemingly accidental consequence of other laws. (See note added in proof, p. 49.)

3.4 Newton's route to an undulatory–corpuscular picture?

I am conjecturing that, despite Newton's speculations in his *Opticks*, concerning the possibility of an all-pervasive medium, and despite the seeming implications, in *Principia*, that he might be happy with an absolute space, Newton carried forward from his youth a deep insight into the relativistic nature of physics. Indeed, he well recognized that his space and his medium, if they 'existed' in any ordinary sense, would have to have profoundly peculiar properties. However, it would be far too much to expect of him that the realities eventually revealed by Maxwell concerning electromagnetism and light, and particularly those revealed by Minkowski and Einstein concerning the essentially four-dimensional structure of spacetime, could have been discernible by him. That would have been a quite unbelievable route for him to follow. Instead, he chose the alternative Galilean-relativistic, but, for the time, equally valid, route. His early relativistic insights, I am suggesting, may well have been largely responsible for an intuitive belief in a corpuscular picture of light. Once he was firmly down that road, and having to come to terms with optical interference phenomena, he was led almost inevitably (and, in a sense, 'correctly') to a picture of light in which corpuscular and undulatory features had somehow to coexist, leading to an eerie foreshadowing of quantum wave–particle duality over two hundred years before its time.

It may well be that there were other 'valid' hints that a combined undulatory and corpuscular picture had to be entertained. Newton's interest in the colours of actual objects, and his alchemical work concerning

the transmutation of substances, must have convinced him of a deep interrelation between light and matter. He certainly speculated that matter and light might be convertible into one another (*Opticks*, *Query* 30). One is reminded of Einstein's early struggles concerning the absorption and emission of light (see Pais, 1980). It was not only Planck's explanation of the black-body spectrum which disturbed Einstein. The whole idea that individual atoms could discretely pick up energy from what, at that scale of things, would have been a barely oscillating and nebulous medium seemed almost preposterous. Einstein's intuitions and insights guided him to the view that, despite the fact that light is waves, light must also be particles: indeed Nature *is* preposterous! It is not inconceivable to me that a few of the many worries that so haunted Einstein in the early twentieth century might already have been vaguely discernible to Newton over two hundred years earlier!

3.5 Quantum mechanics

For the rest of this essay, I shall depart from such speculative considerations of pseudo-history and concentrate on this very preposterous nature of reality that the observational facts of modern quantum physics have presented us with. The questions of quantum reality are closely related to some of those we have just been discussing. We shall find that these questions will lead us into some speculative considerations of a different kind!

Let us remind ourselves of one of the very basic rules of quantum theory, namely *linearity*. If $|\psi\rangle$ and $|\phi\rangle$ are two vectors representing quantum states, then any quantity of the form

$$\lambda|\psi\rangle + \mu|\phi\rangle,$$

where λ and μ are complex numbers not both zero, also represents an allowable quantum state. Recall that multiplying a quantum state-vector by a non-zero complex number yields a state-vector representing physically the same quantum state as before, while all other state-vectors describe physically distinct quantum states. Thus, taking all the different possible ratios λ/μ (including ∞), we are provided with an array of physically allowable alternative states, obtainable by quantum linearity from the original two states. The abstract space of such complex ratios has the topology of a spherical surface – the *Riemann sphere*.

This sphere is implicitly present in any two-state quantum system. Sometimes it has a very clear geometrical role to play; sometimes this role is

rather subtle. It is clearest for the case of the spin-state space of a spin one-half massive particle, say an electron. Here each physically distinct state of spin corresponds to a definite direction in space. This direction is the one with the property that if the spin of the electron is measured in that direction, then the answer to the measurement is: *yes*, with *certainty*. In the case of polarization states of a photon, the geometrical role of this sphere is slightly less direct. Suppose that the photon is moving upwards. Then the north and south poles correspond, respectively, to the states of right-handed and left-handed circular polarization. The points around the equator correspond to the states of linear polarization (Newton's 'sides'), but since a linear polarization state corresponds to an *un*directed line element, the two opposite directions of this line element must correspond to the same point on the equator. More directly, it is the 'square root' Riemann sphere for $(\lambda/\mu)^{1/2}$ which has a clear-cut geometrical correspondence with polarization directions (Stokes vectors).

Sometimes the geometry is not at all direct, such as with the various linear superposition states of two different positions of a particle. Suppose a particle is passing through a pair of slits A and B. If the particle is at A, then we represent this by the north pole of our Riemann sphere. If it is at B, then by the south pole. The various other points of the sphere correspond to states in which the particle is partly at A and partly at B. As we know, it is not a question of mere probabilities of the particle being at A or B, but of *amplitudes* of it being at A or B – whatever that might mean! The space of different possible ways that the particle can be located as it passes through the pair of slits has the structure of a *sphere* – not a line segment, as would be the case with probabilities (the range of probabilities being the real-line segment [0, 1]).

3.6 Physical reality

We are here beginning to be confronted by the odd picture of reality that quantum mechanics presents us with. Do we really *believe* that the particle can be in two places at once, with complex weighting factors? Our intuitions are not so stretched in the case of the spin states of the electron. In that case we can visualize the electron as 'spinning' about an axis which has some definite direction in space. The photon, too, can be thought of as having some definite geometrical polarization state. Some people feel uneasy with such clear-cut geometrical pictures of quantum states. They say that we should not try to form a picture of reality when applying the rules of quantum mechanics: just follow the rules of the quantum mechanical

formalism, do not try to form pictures and do not ask questions about reality! This seems to me to be wholly unreasonable. Physics, after all, constitutes our best way of groping for the true nature of the real world in which we find ourselves. Some would have us believe that physics is to be regarded as merely providing a means of effective prediction about the future. It is true that one of the most powerful shifts in attitude in science came about with the methodology of Galileo and Newton: form a mathematical theory from which predictions can be made and test these predictions against observation. Do not ask *what* it is that constitutes the substance of things, nor *why* things behave as they do. Just ask *how* they behave and try to form an elegant mathematical structure that mimics that behaviour as closely as possible. ('*Hypotheses non fingo*'!)

However, the power of such methodology should not obscure the fact that we are still, nevertheless, groping for the realities of the physical world. The methodology is powerful because it enables us to clear from our minds a good many preconceived and incorrect notions concerning the nature of this reality. It does not, and it should not, lead us to believe that we can dispense with the notion of physical reality altogether – as some would seem to have us believe! What, after all, is the point of making 'predictions about the future' unless we are allowed to assume that there is some reality about this future whose state we are proposing to predict? Despite the operationalistic nature of the methodology of Galileo and Newton, their dynamical theory provided a picture of 'reality' which was very much clearer than that which had gone before. In fact, this picture has become so 'clear' to us that there is now a great reluctance, even after we have been forced to accept the physical existence of quantum phenomena, to believe that 'reality' can take any other form.

It is clear from Newton's *Queries* in his *Opticks*, that he was, indeed, well prepared to speculate on the nature of physical reality and that, in the case of gravity, the existence of a wonderful mathematical theory did not stop him from speculating on 'causes' (cf. his refractive all-pervasive medium referred to above). '*Hypotheses non fingo*' did not apply to *Opticks* (nor, indeed, did it apply to Newton's earlier thinking)! It seems likely, also, that Newton's picture of reality was actually *not* very close to the 'clear' picture of reality that we have subsequently built up from a familiarity with Newtonian dynamics.

3.7 Reality of the state vector

Let us return to quantum mechanics and to the strange pictures that it

presents us with. We have seen that, with electron spin, our geometrical picture of the quantum state-vector is not at all an unreasonable one. However, even here, many would say that we are deluding ourselves, and to form such a picture is misleading. Somehow the electron has only *two* ways in which its spin can point, not a whole spherical continuum of ways. Any experiment that we may perform on the electron to measure its spin can give us but *one bit* of information: the spin can turn out to be one way, or it can turn out to be the other, but it can be nothing in between. That is the way quantum mechanics works.

However, when there are just two such 'alternatives', the space of states is indeed a whole Riemann sphere. We cannot say which the alternatives are. They might be up/down or north/south or east/west or anything in between. There is nothing to choose between any of these pairs of alternatives. But given any *particular* state of spin of the electron, there will be precisely one direction in space for which the state gives certainty that the spin is in that direction. Given the state, there is a 'reality' about the direction in which the state points. An observation *might* be performed measuring the spin in that direction and, if so, the state has to be prepared to say 'yes', with certainty, to that measurement. Somehow the state has to 'know' that direction, even though there is no experiment which can be performed to determine which direction in space it actually is. So long as we are sticking to standard quantum mechanics, the different possible 'realities' for the states of spin of an electron are indeed the points of the Riemann sphere. This is for a system with just *two* alternative states. In a sense, the whole Riemann sphere 'counts' as just two!

For massive particles of higher spin there is also a corresponding geometrical picture for the states. Take the spin to be $n/2$, then in place of the Riemann sphere we have a complex n-dimensional projective space of different possible physical states. But we can still use the Riemann sphere to describe the individual states. Each physical state of spin $n/2$ corresponds uniquely to an unordered set of n points on the sphere of directions.† A characterization of these points is that if the spin is measured in that direction then there is zero probability that the spin turns out to be totally in the opposite direction. Again, the state is uniquely characterized by the fact

† I am told that this result is due originally to Majorana, but I do not know of a reference. The result is not hard to see using 2-spinors (cf. Penrose and Rindler, 1984, p. 162). If $\psi_{AB\ldots N}$ is an n-index symmetric spinor representing the state, then it has a canonical decomposition $\psi_{AB\ldots N} = \alpha_{(A}\beta_B \ldots \nu_{N)}$ (round brackets denoting symmetrization), and the n directions are those represented by the spinors $\alpha_A, \beta_B, \ldots, \nu_N$.

that if measured in any one of these n directions, then with certainty it cannot spin totally the other way. The state has to 'know' that it must respond this way in the event that any one of these measurements is carried out, so, up to proportionality, it is again objectively characterized.

The same objectivity holds, in principle, for any quantum state whatever – so long as the rules of quantum mechanics are presumed to hold rigorously. Suppose the state-vector is $|\psi\rangle$. Then we can consider making an observation corresponding to the observable

$$Q = |\psi\rangle\langle\psi|.$$

The state $|\psi\rangle$ is the only one, up to proportionality, for which the observable Q yields the result 1 with certainty. The state must 'know' that it has to produce this result in the event that the observation Q is actually performed. This is a completely objective property, so the fact that the state is (proportional to) $|\psi\rangle$ is an objective property of the world.

I am expressing a point of view here which, for some reason, is not often maintained. One frequently hears the opposing view that 'the state-vector merely expresses our state of knowledge about a system', or 'the state-vector is expressing a property of ensembles of systems rather than of a single system'. However, the point of view that I am expressing is that the state-vector is clearly defining an objective property of a *single* system. It is not a 'testable' property of that system in the sense that we can perform an experiment on it to determine what its state $|\psi\rangle$ actually is, but it is an 'objective' property of that individual system in the sense that the state of that one system is characterized by the results of experiments that one *might* perform on it.

3.8 Quantum non-locality

Why is this objectivity not something that is more frequently expressed by theorists? The answer to this question is not really clear to me, but I suspect the reason has to do with the strange picture of objective reality that it presents us with. Take the situation considered above in which we had a particle passing through two slits A and B. The objective reality of the system, I claim, actually consists of the particle being partly at A and partly at B, both together in some complex linear combination (since for each choice of complex linear combination of being at A and being at B there is an experiment that could be performed which gives the answer 'yes' with certainty *only* for that particular combination). However, we like to think of particles as tiny point-like objects, not as things that can be in two places at

once. Moreover, we may choose to measure whether the particle is at A, or to measure whether it is at B, and if we find it at one place we cannot find it at the other. Somehow, if we perform such a measurement, we find that at the next moment the particle *is* either at A or B, so we like to think that, a moment before, the particle really *was* in one place or the other, but because of some kind of funny statistics which holds at the quantum level the alternatives must be weighted with complex rather than real coefficients. If we take the realistic view of the state-vector that I maintain, in view of the above discussion, is really almost forced upon us, then we must accept a degree of non-locality: finding the particle at A instantly forces it to be *not* found at B and, conversely, *not* finding it at A instantly forces it to *be* at B. This latter type of observation is an instance of the particularly confusing situation referred to as a 'null measurement'. The mere fact that a detector somewhere has *not* registered the presence of some particle forces the state of that particle to be different from what it was before in some region perhaps distant from the detector itself. It is especially puzzling that the *absence* of a physical interaction can affect the state of a particle in some distant region!

This non-locality is perhaps even more striking when there is more than one particle involved. The state-vector may consist of linear superpositions of different arrangements of the particles. Observing the state of one of the particles may instantly force the others into some state correlated with it, even though they may be well separated from it. As an instance of this sort of thing, let us consider a variant of Bohm's version of the Einstein–Podolsky–Rosen thought experiment. Suppose that we have a particle in a spin-zero state which decays (non-relativistically) into two particles each of spin one-half. Suppose also that there is no orbital contribution to the angular momentum. We can express the various spins in terms of eigenstates of spin in the vertical direction. Then we can think of the initial spin state $|0\rangle$ to be decomposed, after decay, into the difference between two states, one state having one of the particles spinning up and the other down, and the other state having the spins of the two particles both reversed:

$$|0\rangle = |\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle.$$

However, there is nothing special about the choice of the vertical direction here. We could equally well have chosen left/right:

$$|0\rangle = |\leftarrow\rangle|\rightarrow\rangle - |\rightarrow\rangle|\leftarrow\rangle,$$

or, of course, any other pair of opposite directions. Neither of the resulting particles has, by itself, a state. Only the pair considered as a whole has. If we observe the spin of one of these particles to have some particular direction,

then the other instantly acquires an 'objective spin state', namely the state corresponding to spin with certainty in the opposite direction. This applies equally to any direction that we may choose. Moreover, there is no requirement that the two particles be at all close to one another. In principle they could be hundreds of miles apart. By observing the spin of one of the particles, we instantly put the other particle into a spin state whose direction is fixed by our *direction of observation* on the first particle! (We note that no message can be sent from one particle to the other by this means.) This is a decidedly non-local picture of reality.

I am regarding the total state on the right-hand side of these equations as being 'real'. But how else are we to regard it? We are always at liberty to choose to measure the spins of neither particle but, instead, to reflect both of them – carefully, so as not to disturb their spins – back to their original position so that they can recombine as a spin-zero state. The spin of the recombined state must indeed be zero with certainty, and the possibility of observation of this final spin assigns a 'reality' to the non-local spin state of the combined system. Other situations of this same general nature exhibit a similar non-locality. This is particularly striking when the two emitted particles are photons since the effect of observing one of them would, if it were a signal in the ordinary sense, have to travel faster than light in order to reach the other one. The experiments of Freedman and Clauser (1972) and Aspect (1976) and coworkers, etc., have shown that a non-locality of this type (over a distance of some twelve metres, in the latter case) is an actual feature of the world we live in and not just a theoretical fiction.

3.9 Quantum mechanics and macroscopic physics

We see that this quantum view of reality is very different from the one that we have become accustomed to from classical physics, where particles can be only in one place at a time, where the physics is local (except for action at a distance) and where each particle is a separate individual object which, when it is in free flight, can be considered in isolation from any other particle. All these classical conceptions must be overturned once we accept the reality of the state-vector. It is perhaps little wonder that most people are reluctant to do this. However, the quantum view of reality is more worrisome even than this. For the linear superpositions between different particle positions can get coupled in with macroscopically differing states of *macroscopic* objects. In that case, according to the strict rules of quantum mechanics (and assuming that no 'observation' is deemed to have been made) we must carry over the complex linear superpositions to the

macroscopic objects themselves. Thus we can be presented with a situation like that of Schrödinger's cat, where we seem to have to consider states representing linear combinations of a dead cat and a live cat, with complex coefficients. There is a very reasonable reluctance to accept such a state as actually describing reality! Surely, at that level the cat *is* either dead or alive. However, there still remains the possibility, in principle, that one might perform the observation described by $Q = |\psi\rangle\langle\psi|$, as suggested above, that requires objectivity for the state $|\psi\rangle$ *whatever* it is. In this instance, where $|\psi\rangle$ represents some complex linear combination of a dead cat and a live one, it might well be objected that the observation described by Q is quite out of the question practically. But quantum theory, as it stands, has nothing to say about what is 'practical' or 'impractical'. To place some restriction on which operators Q are allowed as observables, and which are not, would be to introduce a *change* into the theory.

Some people take the view that a quantum linear superposition at that level is still perfectly in order. A conscious human observer could also be in such a superposition but, because of some quirk of the nature of conscious perception, his consciousness would 'split', so that each state of awareness would perceive just one of the macroscopically distinct states under superposition, and would not be aware of the actual superposition. This (or something like it) is the many-worlds view of Everett and his followers. Personally, I should be most reluctant to go to such lengths to save the superposition principle for macroscopic bodies. Not least, it is worrisome to have to rely on a yet undiscovered theory of consciousness in order to find agreement between the picture of things presented by theory and what is observed in practice. It seems to me that something must go wrong with the superposition principle a good deal before this. Macroscopic physics *is* like classical physics, where particles occupy single localized places and behave as individual separate entities.

3.10 Linearity and time-evolution

So far, I have not discussed the time-evolution of quantum states. As we know, there appear to be two completely different modes of evolution in quantum theory – and this decidedly odd fact has, no doubt, added to the reluctance that many people may feel about taking the state-vector as actually representing reality. On the one hand, we have unitary evolution (Schrödinger or Heisenberg, according to one's description). This is completely deterministic, as in classical physics, and it preserves quantum linearity. On the other hand, we have 'state-vector reduction', which is

probabilistic and does not preserve quantum linearity. If, as I am proposing here, one takes a realistic view of the state-vector (and supposing that one rejects the many-worlds view – or even if one does not), then one must come to terms with these two different modes of evolution.

Some would take the line that state-vector reduction is an illusion which appears to arise because one is ignoring the complicated interconnections between the system under consideration and the outside environment. The claim is, on this view, that unitary evolution *would* hold if we could take everything into account, but, since we are ignoring the outside, we are led to an approximation which appears like a reduction process. I have to say that I do not find this viewpoint at all plausible. I really do not see how the linear superposition of Schrödinger's dead and alive cat can get resolved to one alternative or the other, simply by taking the environment into account. The best that one can hope to do with this type of procedure is to show that somehow the cat can be treated *in practice* as some probability-weighted mixture of alternatives. It cannot explain how the reality of the situation can shift from a complex linear superposition to alternatives. Unless we change the theoretical structure of quantum mechanics at some stage we still have the possibility *in principle* of performing the measurement Q referred to above, so 'reality' must remain the complex linear superposition!

Another suggestion is what is known as the 'Wigner view' (Wigner, 1961), according to which unitary evolution holds until consciousness gets involved in the physical system. At that stage, some non-unitary evolution takes over and state-vector reduction objectively takes place. Again I am worried that a theory of consciousness seems to be required in order that the theory can be used. But a stronger objection, to my mind, is that the view seems to assign a privileged role to those corners of the universe where consciousness resides – and they may be very rare indeed. The objective evolution of the state-vector would be quite different in those corners from elsewhere. The effect might not be 'observable' but the asymmetry in the pattern of time-evolution presents a disturbing view of reality.

My own picture is different from any of these, though it shares with the Wigner view the idea that unitary evolution is to be supplanted, under certain physical circumstances, by an objective evolution which closely approximates state-vector reduction. However, I am *not* suggesting that consciousness is needed for this. The reduction process would normally be completed long before the results of the reduction are perceived by any conscious observer. Indeed, an actual conscious observer would normally be playing no role whatever in the reduction process. Instead, I am

suggesting that, at a certain level of scale, a failure of quantum linear superposition becomes important, and unitary evolution is taken over by some specific non-unitary procedure. It might be that this procedure can be described by some non-linear (and non-local) differential equation, non-linear instabilities arising at a certain scale, so that the system then 'flops' into one or the other of two macroscopically distinguishable states. On the other hand, the procedure might have to be described in some quite new way. Non-linearity is essential, since the *conventional* reduction procedure does not preserve quantum linearity; but it is not clear to me that a differential equation is the right thing to look for. Also I do not wish to prejudice the issue as to whether the reduction procedure should be genuinely stochastic or whether there should be some underlying determinism which just has the *appearance* of randomness. The linearity of quantum mechanics would, on this view, be merely an approximation, albeit an excellent one at the scales where quantum mechanics has been tested.

Many people feel uncomfortable about dispensing with quantum linearity, rightly pointing out that complex linearity is the source of so much of the extraordinary mathematical elegance of conventional quantum theory. To this I can only respond that we have a precedent. The remarkable mathematical elegance of Newton's theory of universal gravitation, as so magnificently expounded in *Principia*, also owes much to the fact that the total gravitational force at any point is a *linear* function of the sources. Yet when this theory was finally superseded by a more comprehensive one, namely Einstein's general theory of relativity, it was by an astonishing *non*-linear theory of even greater mathematical elegance! I would anticipate that the appropriate theory which replaces our present quantum ideas must, when it comes, exceed even quantum mechanics in its mathematical beauty, economy and grandeur. Unfortunately, I am in no position to make clear suggestions as to what form such a theory might take. However, I *am* making a definite proposal as to the kind of physics which should, in my view, be relevant. This proposal leads us to a fairly concrete suggestion, enabling us to give, in principle, very rough estimates as to when linear evolution should fail and be taken over by some non-linear procedure resembling state-vector reduction.

3.11 Quantum gravity and time-asymmetry

The new physical ingredient which must be added to quantum theory is, I claim, *general relativity*. The new physical theory we are seeking is thus a form of *quantum gravity*. As of now, no satisfactory theory of quantum

gravity exists. Ideas abound and vast calculations continue to be performed, but there is still no adequate theory. My own ideas concerning the form that this theory should take – and even concerning the very purpose of such a theory – seem to differ from those of most workers in the field. Most people seem not to be concerned with trying to alter the structure of quantum mechanics. On the contrary, the impulsive reaction has generally been to modify the structure of general relativity instead (to supergravity, more complicated Lagrangians, higher dimensions, or strings), as soon as the encounter between these two great theories runs into difficulties!

Why do I believe that quantum theory is likely to require revision when unified with general relativity? Aspects of my views on this question have appeared in other articles (Penrose, 1976, 1983, 1986); also, I refer the reader to other related viewpoints which had been put forward earlier (Károlyházy, 1966; Károlyházy, Frenkel and Lukás, 1986; Komar, 1969). Part of the reason for believing that the true quantum gravity theory may entail a change in the structure of quantum mechanics is to be found in the very difficulty that one encounters in trying to 'quantize' general relativity in a standard way. Often these difficulties seem to confront one with foundational problems concerning quantum observations. I do not propose to go into a discussion of these here. Many of these arguments are familiar in the literature (cf. Wheeler, 1964; Kuchař, 1981). In my opinion there is a considerable force behind these arguments. Attempts to build up an *S*-matrix theory of gravity from perturbation series starting from flat spacetime, for example, simply do not face up to many of these long-standing foundational problems.

I wish to stress, on the other hand, a point of a different nature – one which might even be called an *observational fact* of quantum gravity! One of the few situations in nature where it is generally accepted that quantum gravity effects should be important is in governing the physics at *spacetime singularities* (or, at whatever quantum-geometrical regions are to replace such classical singularities). From the assumption that cosmic censorship holds, we can deduce that these singularities can be divided into two classes: past singularities (out of which particles can emerge) and future singularities (into which particles can disappear). (See Penrose, 1978.) The one past singularity that we know of is the *big bang*. Future singularities are expected to arise inside black holes and perhaps in the big crunch (if the universe eventually recollapses to a singularity). Now the big bang was very far from being a 'generic' singularity. We know this both from the observed uniformity of the early universe and from the very existence of a second law

of thermodynamics (Penrose, 1979). (This argument is not affected by current considerations of 'inflation'.) If we imagine the big bang time-reversed, we see more clearly how special it really is. In a generic big crunch, we expect many black holes congealing finally into one very complicated final singularity. The geometry close to the end would be extremely irregular and not at all close to the highly uniform and ordered singularity that is observed to be the case with the big bang. In the final big crunch (if it occurs – or, in any case, as a theoretical consideration) we have an irregular, extremely high entropy singularity. To a somewhat lesser (but still very considerable) degree, the situation is similar for the singularities inside black holes; each such singularity would be of a very high entropy type. However, the big bang was a singularity of a uniform and very low entropy type. The lowness of this entropy lay in the fact that the gravitational degrees of freedom were somehow initially not 'thermalized' with the matter and, moreover, have not much become so because of the weakness of the gravitational coupling. As time progressed, the matter in the universe has gradually been able to take advantage of this imbalance and clump gravitationally. Our Sun, itself, is a result of such gravitational clumping and has become the hot spot in the sky that all life on Earth ultimately relies upon as its source of low entropy.

Using the Bekenstein–Hawking formula for black-hole entropy, we can estimate the entropy that a final big crunch would have, and hence deduce how extraordinarily special the big bang actually was. The answer is remarkable: for a closed universe with, say, 10^{80} baryons, we find that the big bang was special to the degree roughly

$$\text{one part in } 10^{10^{123}}$$

where this figure refers to the region of phase-space volume that corresponds to a big-bang singularity of the uniformity that is observed in the universe as we know it. For a larger number of baryons, this figure would be even larger. Allowing for inflation makes virtually no difference to the argument.

The specialness of the big bang can be accounted for if we assume that for some reason – presumably as a consequence of the quantum gravity theory that we are striving for – the Weyl curvature is necessarily zero (or at least very small compared with the Ricci curvature) at any past singularity. The same constraint cannot apply to future singularities, for, if it did, we should not have a second law of thermodynamics. Moreover, there would then be all sorts of teleological behaviour preventing the formation of black holes. We see that this one 'observed' consequence of quantum gravity is grossly

time-asymmetrical! From this we can deduce the important moral: *the true quantum-gravity must be a time-asymmetrical theory.*

3.12 Time-asymmetry of state-vector reduction

Notice that this is not at all the expected conclusion if we think of quantum gravity as arising as the result of the application of some standard quantum-mechanical procedure (which would be time-symmetrical) to standard general relativity (which *is* time-symmetrical). What is needed is some new procedure which allows for some time-asymmetry.

Now, state-vector reduction – or the more complete process that underlies it – is, I claim, a time-*a*symmetrical process. In the first place, the realistic view of the state-vector that I have been promoting itself entails a time-asymmetry that is not necessarily present in other attitudes to quantum mechanics. For we note that immediately *after* an observation is made, the state-vector is in an eigenstate of the relevant Hermitian operator, but this does not generally apply to the state immediately *before* an observation. (Time-symmetry for quantum evolution can be partially reinstated at the expense of the realistic view, where one regards transition probabilities as providing the only observational aspects of quantum theory; see, for example, Penrose, 1979, p. 584. However, I am now disputing my earlier acquiescence in the common contention that transition probabilities are time-reversible – as we shall see in a moment!) Note that this picture is tied up with the fact that our realistic view seemed forced upon us because of our presumed freedom, in principle, to carry out a quantum-mechanical measurement corresponding to the Hermitian operator of our choice, here the operator $Q = |\psi\rangle\langle\psi|$. The state has to be 'prepared for' the possibility that this measurement might 'at any moment' be performed on it. The use of such an argument in the past direction of time seems to carry much less weight, since one does not seem to have the freedom to carry out 'unexpected' experiments in the past! One thinks intuitively that the state 'already knew' what was done to it in the past, but 'cannot know' what we might unexpectedly do to it in the future. Of course these feelings may be entirely inappropriate and illusory. For very minute intervals of time, I would myself not give them much credence. Yet, for a state which has persisted for a while, I cannot help feeling that there is something in the 'common-sense' attitude that the state is being governed by what was done to it in the past rather than by what is going to be done to it in the future. (This use of 'governed' has not really much to do with the determinacy/indeterminacy question. I am referring, here, to the requirement that the

state be an eigenstate of the past measurement immediately after it was performed, but not an eigenstate of the future measurement immediately before if will be performed.) For the moment, at least, I am allowing this use of 'common sense' to instruct us that our objective state-vector must indeed behave in a time-asymmetrical way under the reduction process.

This time-asymmetry seems to be bound up in some way with our 'freedom of choice' to perform measurements. It is probably also, therefore, bound up with the second law of thermodynamics – although the precise relation between this question of 'choice' and the second law is, to my mind, somewhat obscure. However, there is a second aspect of time-asymmetry in the 'measurement process' which, in a certain way, seems also to be connected with the second law. Imagine a two-state system, with an orthogonal basis for states that I symbolically write as $|\uparrow\rangle$, $|\downarrow\rangle$. Let us assume that the system is initially prepared in the state $|\uparrow\rangle$. Suppose that a measurement is then performed on the system, the eigenstates of the measurement being $|\rightarrow\rangle$ and $|\leftarrow\rangle$. (This need have nothing to do with spin. I just mean these to be orthogonal states different from the original pair.) There will be some resulting probability that the outcome is $|\rightarrow\rangle$ and some probability that it is $|\leftarrow\rangle$. Two possible universe histories that must be allowed by the reduction procedure are therefore:

first $|\uparrow\rangle$ and then $|\rightarrow\rangle$

and

first $|\uparrow\rangle$ and then $|\leftarrow\rangle$.

Now if this procedure were time-symmetric, it would have also to allow both

first $|\uparrow\rangle$ and then $|\rightarrow\rangle$

and

first $|\downarrow\rangle$ and then $|\rightarrow\rangle$,

if we assume that the final state is given as $|\rightarrow\rangle$, and also both

first $|\uparrow\rangle$ and then $|\leftarrow\rangle$

and

first $|\downarrow\rangle$ and then $|\leftarrow\rangle$

for a given final state $|\leftarrow\rangle$ (where the words 'first' and 'then' are being used in their normal temporal sense), the same probabilities arising as before (in reverse order), where now it is $|\rightarrow\rangle$, or $|\leftarrow\rangle$, that is, 'given' in the time-reversed reduction procedure.

Now, in practically any plausible experimental set-up, the input $|\uparrow\rangle$, and the two alternative outcomes $|\rightarrow\rangle$ and $|\leftarrow\rangle$, are all perfectly reasonable

things to occur. However, $|\downarrow\rangle$ will normally be 'absurd' – at least in the sense that it represents an extreme improbability which violates the second law – this probability being nothing like what would be predicted by the standard quantum-mechanical computation. For example, one might have a Stern–Gerlach apparatus preparing spin one-half atoms in a spin-up state ($|\uparrow\rangle$). Interposed in the beam we insert a second Stern–Gerlach apparatus oriented at right angles so as to measure the spins in the left/right direction ($|\rightarrow\rangle$, $|\leftarrow\rangle$), and the beam duly splits as required. However, if we try to imagine the alternative input, corresponding to $|\downarrow\rangle$, we find that, instead of being the alternative beam of the original Stern–Gerlach apparatus, it is yet another beam, coming from the laboratory wall or some irrelevant part of the apparatus! Such a behaviour would normally be thought of as extremely improbable, if not totally impossible – and not *at all* that given by the quantum probabilities applied in the time-reversed sense.

An even more elementary situation of this type occurs if we consider a photon, coming from a laboratory source (state $|\uparrow\rangle$), and simply reflect it off a half-silvered mirror, where we set up photo-cells in both the reflected and transmitted positions (states $|\rightarrow\rangle$ and $|\leftarrow\rangle$, respectively). The state $|\downarrow\rangle$ represents the photon coming from such a direction that if reflected it would reach the transmitted position and if transmitted it would reach the reflected position. There would normally be no source in that direction. The photon, in state $|\downarrow\rangle$, like the atoms above, would have to be just 'absurdly' ejected from the laboratory wall!

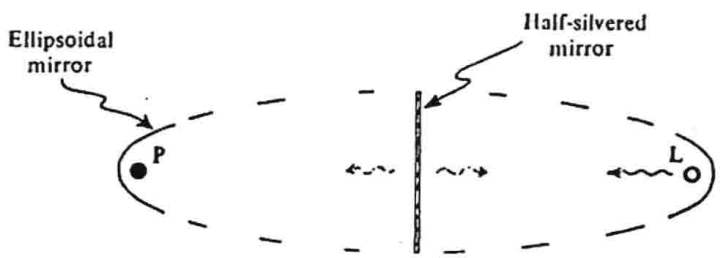
I believe this to be a key issue; and, since it is easily misunderstood, let me try to be somewhat clearer about the sense in which these probabilities are to be interpreted. It is necessary to adopt a viewpoint on this which allows conditional probabilities to be understood in a way that is totally unbiassed with respect to the direction in which time is taken to be running. Let me simplify the above experimental set-up even further so that we have just one photo-cell and one 'lamp' (the photon source). Half-way between the photo-cell and lamp we have a half-reflecting mirror whose plane is perpendicular to the line joining photo-cell to lamp. If desired, we can add an ellipsoidal mirror (a prolate ellipsoid of revolution) surrounding all three, the photo-cell **P** being at one focus and the lamp **L** being at the other, so that we need not worry about the direction in which the photon enters or leaves the photo-cell or lamp. (See Fig. 3.1.) There are four possible ways that the photon can go. It can start at **L** and end at **P**, or it can start at **L** and end at **L**. (These are the 'normal' cases, of everyday experience.) Moreover, it can start at **P** and end at **P**, or it can start at **P** and end at **L**. (These cases are also

needed in order to complete the space of quantum-mechanical transitions.) The respective spacetime descriptions of these four possibilities are given in Fig. 3.2.

Now suppose that this 'experiment' is performed a very large number of times at various different locations throughout spacetime (where, to be on the safe side, I am restricting these locations to be in our own galaxy in the present era). To see what the various conditional probabilities are, we examine the spacetime (Fig. 3.3) and select those cases in which the in-state is the one desired ($|..\rangle$) and count the various alternative out-states ($\langle..|$) (normal use of conditional probabilities), or else we select those cases in which the out-state is the one desired ($\langle..|$) and count the various alternative in-states ($|..\rangle$) (time-reversed use of conditional probabilities). All the quantum-mechanical matrix elements $\langle \mathbf{P}|\mathbf{L}\rangle, \langle \mathbf{L}|\mathbf{L}\rangle, \langle \mathbf{P}|\mathbf{P}\rangle, \langle \mathbf{L}|\mathbf{P}\rangle$ have squared modulus equal to one-half. These give the *correct* probabilities if we select for the in-state $|\mathbf{L}\rangle$ and count how many times the photon ends up in $\langle \mathbf{P}|$ and how many times it ends up in $\langle \mathbf{L}|$ (closely 50 per cent of each). This is using the quantum-mechanical probabilities in the normal direction in time. However, if we select for the out-state $\langle \mathbf{L}|$ and count how many times the photon started out in $|\mathbf{P}\rangle$ and how many times it started out in $|\mathbf{L}\rangle$ (closely 0 per cent and 100 per cent, respectively), we find that the reversed-direction quantum-mechanical calculation gives completely the *wrong* answer (50 per cent for each)!

It is no good trying to retrieve the situation by bringing in other factors that we do or do not know about the universe (e.g. the second law of thermodynamics). We do not ask that other factors be taken into consideration when calculating quantum-mechanical probabilities in the *normal* direction in time. Quantum-mechanical probabilities are supposed

Fig. 3.1. Is state-vector reduction time-symmetric? In this simple experimental set-up a photo-cell P and a lamp L are separated by a half-silvered mirror. A photon travels between them.

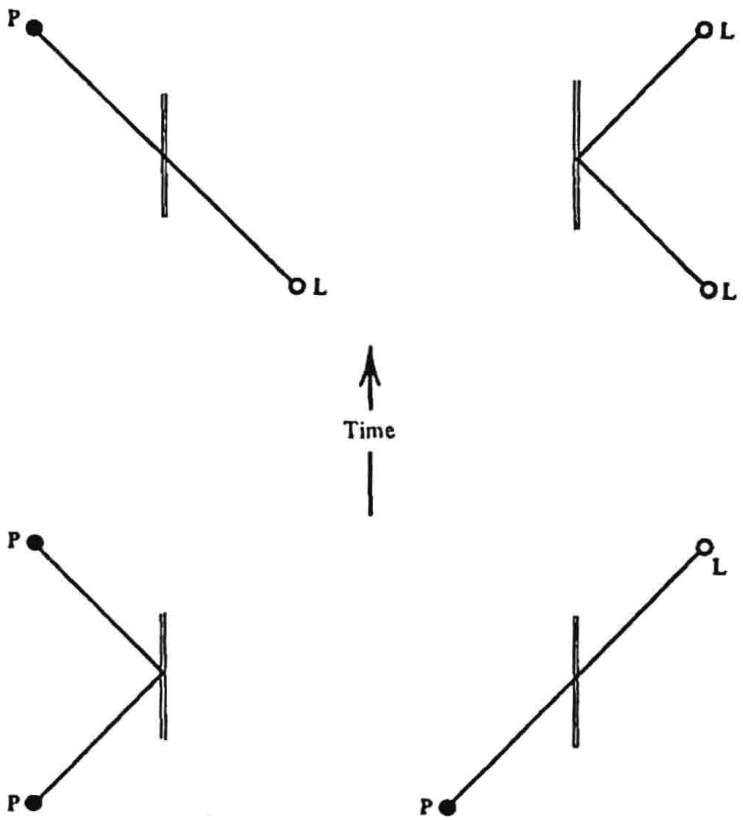


 Newton, quantum theory and reality. *Three hundred years of gravitation.* Eds S. W. Hawking and W. Israel, 17–49. 1987

to be entirely stochastic, and not influenced by any other such factors. The standard rules of quantum mechanics were obtained by observing (in our own galaxy in the present era!) the way in which the probabilities behave in the normal direction in time. These particular quantum-mechanical rules for calculating probabilities simply *do not work* when used in the reverse direction in time.

I am proposing to disallow these 'absurd' states (like the in-state $|P\rangle$ above), and simply *accept* that the objective physical process underlying state-vector reduction must be time-asymmetrical. (Most such 'absurd' states, like the in-state $|P\rangle$, would, when traced backwards in time, ultimately becomes incompatible with the hypothesis of vanishing Weyl

Fig. 3.2. Spacetime diagrams of the four different ways that the photon can travel. The top two diagrams depict the 'normal' situation, with the photon coming from the lamp. The two lower possibilities are needed in order to complete the list of possible transitions between the quantum states.

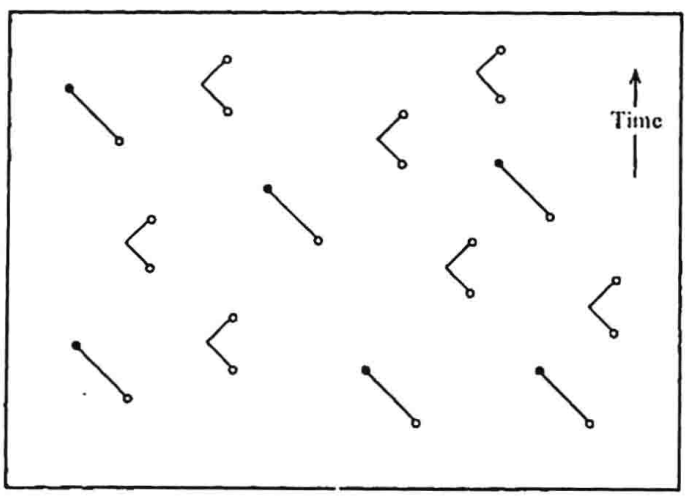


curvature at initial singularities, since they are in contradiction with the second law of thermodynamics.) Thus, one input can have several alternative possible outcomes, but there would not generally be two different inputs leading to the same outcome. This would mean that if we follow a path in the phase space of an isolated system, then at certain stages, namely where the state-vector reduction process objectively takes place, this path would split into several paths, but different paths would not generally unite together in this process. Thus, the evolution in phase space that is implied by such a state-vector reduction process will be time-asymmetric: paths can sometimes bifurcate but they do not rejoin.

3.13 Reduction and the (longitudinal) graviton count

Apart from merely this fact of time-asymmetry that I am claiming for the reduction process, is there any more direct evidence to link this process with gravity? I have argued elsewhere that there may well be (Penrose, 1979, 1981, 1986). The quantum gravity effect that I claim must restrict past singularities to have vanishing (or small) Weyl curvature also *disallows white holes* (the time-reverses of black holes). For an isolated system of large enough mass, this entails that paths in the phase space must sometimes become *united* into a single path. This occurs because information can get

Fig. 3.3. The entire space-time is examined. If the *initial* state is selected for, the final state ratios are compatible with the quantum-mechanical calculation, but if the *final* state is selected for, the same calculation gives quite the *wrong* answer for the initial state ratios.



 Newton, quantum theory and reality. *Three hundred years of gravitation.* Eds S. W. Hawking and W. Israel, 17–49. 1987

lost at the singularities, all of which are of the *future* type (into which matter can disappear: black holes) rather than of the past type (from which matter could have emerged: white holes), the latter having to be effectively excluded owing to the above-mentioned claimed quantum gravity effect. As I have suggested elsewhere (Penrose, 1981, 1986), it is very plausible to suggest that the uniting of phase-space paths that result from the above singularity discussion is exactly balanced by the splitting of phase-space paths which I argued earlier should result from the proposed objective state-vector reduction procedure. I am not, of course, arguing for any direct connection between specific spacetime singularities and specific instances of the reduction process. Clearly one does not need a black hole in the laboratory in order to perform a Stern–Gerlach experiment! I am speaking merely of an overall balance between phase-space loss from spacetime singularities and phase-space gain arising from state-vector reduction. The argument is that if the first is a quantum gravity effect, then so also must be the second.

In earlier accounts I have phrased things in terms of some (undefined) concept of 'gravitational entropy'. My present opinion is that this concept is both inappropriate and unnecessary for the discussion. (I am grateful to Boris Zel'dovich and Bob Wald for some very pertinent criticisms of my earlier view.) Instead, I believe what is required is some measure of (longitudinal) graviton number. The idea is that when two states in linear superposition become differently coupled to the gravitational field, to the extent that the difference between the fields is of the general order of *one (longitudinal) graviton*, then failure of linear superposition becomes significant, and some non-linear procedure (instability?) forces the state out of the superposition and into one alternative or the other. This procedure, if it could be found, would provide our sought-for objective state-vector reduction.

The immediate question that arises is whether the orders of magnitude involved make this seem at all plausible. While I have not been able to explore this question in a great deal of detail, it seems to me that the signs are not altogether discouraging. Let me give an example, which I shall try to describe according to the kind of picture that I have in mind. Suppose we have an atom surrounded, at some distance, by a particle detector – say a cloud chamber or a bubble chamber. The atom decays, emitting a charged particle. As the particle moves away from the atom, its *reality* consists of a spherical wave moving outwards and centred on the atom. Quantum linearity holds excellently at this level, and the wave can, if preferred, be viewed as a linear superposition of straight particle tracks directed outwards

from the atom. As these tracks enter the detector, they each cause streaks of ionization which in turn cause trails of droplets, or bubbles, to be formed. The reality of the situation is still a complex linear combination of all these possible different trails together, so long as the movements of energy are sufficiently small that the coupling to gravity is significantly below the level of one graviton. Eventually, however, when the droplets, or bubbles, grow to a sufficient size, the differences between the gravitational fields produced by the different linearly superposed trails will grow to become quantum-mechanically significant – by which I mean that the (longitudinal) graviton count for the differences between the various gravitational fields of the individual trails reaches order unity. When this happens, non-linear effects become important and the system 'flops' into a state in which the spacetime geometry is (at the level of one graviton or so) well defined. Just one of the complex-linearly superposed spacetime geometries is singled out, and so just one of the possible trails reaches realization. The idea is that, in a sense, Nature abhors complex linear combinations of differing spacetime geometries!

One must do a calculation to see if this is in any way plausible. My own crude attempts at this have been superseded by a calculation performed by Abhay Ashtekar. Let us, for simplicity, assume that we have just a single droplet. (The case of a bubble is essentially similar: we simply count its mass as negative – which occurs in the formula *squared*. This mass is the difference between the mass of the bubble and that of the ambient medium.) Let the droplet's mass be m and its radius be r. Let the radius of the material from which the droplet has condensed be R. Take everything to be spherically symmetrical, and suppose, crudely, that there is a region of vacuum formed between the radii R and r. Adapting, to the case of (linearized) gravity, Zel'dovich's (1966) procedure for estimating the number of photons in a classical free electromagnetic field and applying the expression to the situation of a static field with source (where the procedure does not strictly apply – but what else can one do?), and adopting a particular gauge condition, Ashtekar obtains the following expression for the expectation value of the number of (longitudinal) gravitons arising in the region between radii r and R (this being the region where the linearized Weyl curvature is non-zero):

$$768\pi^3(m/m_{\rm P})^2 \log(R/r).$$

Here $m_{\rm P}$ is the Planck mass ($\sim 10^{-5}$ g). The coefficient $768\pi^3$ is about twenty-four thousand, so effects at the 'one-graviton' level could show up

when m is around 10^{-7} g or so. The hope is that the corresponding formula for a whole string of correlated droplets (or bubbles), and other cases without spherical symmetry, would lead to a somewhat smaller mass giving a one-graviton effect.

The situations with other types of measurement, such as (developed?) photographic plates, spark chambers, computer memories and, most importantly, bundles of nerve signals (since the retina can be used as a quantum detector for small numbers of photons – and because the only time we *know* that reduction has taken place is when we personally experience it!) remain to be examined. My impression is that it would be more encouraging if plausible theoretical reasons could be found to suggest that the non-linear reduction process should take effect at a somewhat smaller scale than is estimated at present. Perhaps taking into account asymmetries, or detailed granular structure, might lead to such a conclusion. Alternatively, it might be that one needs a slightly different criterion for the onset of reduction, still involving a 'graviton' criterion, but where there would be significant probabilities for reduction over short timescales even well below the 'one-graviton' level. I feel that things are not too discouraging, in view of the primitive nature of all these considerations, so far.

3.14 Non-locality in quantum geometry

From the discussion given above, any physical reduction process must have decidedly non-local and peculiar features. In this respect, it is encouraging that the expression for graviton number in a linearized gravitational field is indeed a non-local one. However, the reader should be warned, in my opinion, that any successful objective physical theory of state-vector reduction must be non-local in a way that fundamentally affects even the very fabric of spacetime. I do not expect, except perhaps at some temporary provisional level, that we shall be able to get away with a theory which describes an objective reality taking place *within* some ambient spacetime. The spacetime must itself become subject to this non-local description.

One reason for believing this stems from consideration of the kind of non-locality that arises in situations of the Einstein–Podolsky–Rosen type as exhibited by the experiments of Freedman and Clauser, and Aspect and his coworkers. We can obtain a consistent description of the photon pairs using the conventional quantum-mechanical reduction procedure, but this must be applied *simultaneously* over a region of space encompassing both photons: a partial measurement is made on the system when one photon is

observed, and this instantly affects the state of the other photon, putting that one into a state which becomes subject to the observation on *it* which immediately follows. However, this picture is not at all Lorentz invariant. If we allow a different Lorentz frame to describe the situation, we can arrange that it is the observation on this second photon which occurs first and causes the partial reduction, the result of which 'then' becomes subject to the original observation on the first photon. In this second Lorentz frame our picture of 'reality' is completely different from the one in the first frame – where I am taking the conventional description in terms of evolving and reducing state-vectors to represent 'reality'.

At this point we could simply abandon relativity for our picture of reality, if we wish. There would be no actual conflict with the results of experiment if we do so. But there is conflict with the *spirit* of relativity in such a procedure (cf. Bell, 1987). Once we adopt a non-relativistic view for our picture of reality, we find it to be a 'fluke' that relativity actually holds. To adopt such a view would be to ignore the insight that Einstein impressed upon us in 1905 – and also the closely related insight that I am suggesting guided the youthful Newton to an important aspect of his views on the nature of light. This is the insight that relativity should be taken as a *principle* rather than as an accidental feature of the precise form of the physical laws. It may be argued that cosmology (and, indeed, the big-bang singularity) now supplies us with a rest-frame, so the case is certainly debatable. Nevertheless, my own money, for what it is worth, would go on a description of reality which, at this level, should indeed be relativistically invariant. The best suggestion that I can make at this stage would be for a picture involving some sort of partially formed, partially bifurcating spacetime, where the nature of the spacetime has not been adequately resolved until the second photon observation has taken place. I have been suggesting, after all, that the reduction process should be an intimate feature of the interrelation between quantum mechanics and general relativity. General relativity describes the structure of spacetime, so, when quantum mechanics becomes intimately involved, we must expect some radically different pre-spacetime notions to enter the picture.

Another puzzle concerning the reduction procedure is the one mentioned above concerning 'null measurements'. *Not* observing a particle in a detector can effect a partial reduction of the state-vector. On the view that I am putting forward, this is not unreasonable. The spacetime geometry which gets coupled to the situation resulting from the particle entering the detector differs from the geometry resulting from it not entering the detector;

and their difference might well reach the one-graviton level. At this point Nature would have to choose between (partial?) spacetime geometries. She might well choose the geometry in which the particle has not entered the detector – and that choice would non-locally affect the state elsewhere.

How is one to develop such a non-local quantum spacetime theory? I have my own ideas about what might be the fruitful line to pursue, but it would not be appropriate for me to try to enter into these here since the ideas, even after very many years of labour, are still much too ill-formulated to suggest any reliable insights. It will surprise no-one if I say that it is still my opinion that the formalism of twistor theory, with its essentially non-local description of spacetime ideas, ought to provide an important input to such a quantum spacetime theory (cf. Penrose and Rindler, 1986). Also, I still have some hankering after ideas that I entertained many years ago in evolving a theory of spin-networks (Penrose, 1971, 1972). Thought experiments of the Bohm–Einstein–Podolsky–Rosen type had an important motivational role to play in that theory, and the idea was to build up the concept of space as a limiting implicit structure when large numbers of particles are involved. However, neither twistor theory nor spin-network theory have any time-asymmetrical ingredients, as things stand. It is clear to me that some essential new input is needed.

Whatever the future may hold for the development of physical theory, it is clear that quantum theory, relativity, and even Maxwell's electromagnetic theory, have led us – or should have led us, if we have examined the facts – to a view of reality very different from that presented by classical Newtonian particle mechanics. From his writings, it is equally clear that Isaac Newton was not, in the strict sense, a 'classical Newtonian'. His view of Nature contained very much more mystery than that which the standard picture of Newtonian mechanics now conjures up. Perhaps his readiness to put forward such a strange corpuscular yet wave-like conception of light is indicative of his belief in the depths of Nature's mysteries. If so, he was certainly right!

Acknowledgements

I am grateful to Abhay Ashtekar and Dipankar Home for valuable suggestions, and to Karel Kuchař for very helpful criticisms of the manuscript. I thank, also, the Institute for Theoretical Physics at the University of California in Santa Barbara for its hospitality and stimulation, where discussions with D. Boulware, I. Bialynicki-Birula, E. T. Newman, W. G. Unruh and R. Wald were most valuable. This research was supported

in part by the National Science Foundation under Grant No. PHY82-17853, supplemented by funds from the National Aeronautics and Space Administration, at the University of California at Santa Barbara and in part by the U.S. Air Force Office of Scientific Research.

References

Arnol'd, V. I. (1978). *Mathematical Methods of Classical Mechanics*. Springer: New York.
Aspect, A. (1976). *Phys. Rev.* **D14**, 1944.
Bell, J. (1987). Beables for Quantum Field Theory in *Quantum Implications: essays in honour of D. Bohm* (eds. B. J. Hiley & S. D. Peat, Routledge & Kegan Paul) p. 227
Drake, S. (1957). *Discoveries and Opinions of Galileo*. Doubleday: New York.
Freedman, S. J. and Clauser, J. F. (1972). *Phys. Rev. Lett.*, **28**, 938.
Galilei, G. (1638). *Dialogues Concerning Two New Sciences* (Macmillan edn, 1914). Dover: New York.
Károlyházy, F. (1966). *Nuovo Cim.*, **A42**, 390.
Károlyházy, F., Frenkel, A. and Lukás, B. (1986). On the possible role of gravity in the reduction of the wave function. In *Quantum Concepts in Space and Time*, ed. C. J. Isham and R. Penrose. Oxford University Press: Oxford.
Komar, A. B. (1969). *Int. J. Theor. Phys.*, **2**, 257.
Kuchař, K. (1981). Canonical Methods of Quantization in *Quantum Gravity 2: a second Oxford Symposium* (ed. C. J. Isham, R. Penrose & D. W. Sciama, Clarendon Press, Oxford) p. 329
Mache, E. (1972). *The Science of Mechanics*. Open Court: La Salle, Ill.
Newton, I. (1730). *Opticks*. (1952) Dover: New York.
Newton, I. (1687). *Principia*.
Pais, A. (1980). *Subtle is the Lord*. Oxford University Press: Oxford.
Penrose, R. (1971). Angular momentum: an approach to combinatorial space-time. In *Quantum Theory and Beyond*, ed. T. Bastin. Cambridge University Press: Cambridge.
Penrose, R. (1972). On the nature of quantum gravity. In *Magic Without Magic*, ed. J. R. Klauder. Freeman: San Francisco.
Penrose, R. (1976). *Gen. Rel. Grav.*, **7**, 31–52.
Penrose, R. (1978). Singularities of space-time. In *Theoretical Principles in Astrophysics and Relativity*, ed. N. R. Leibowitz, W. H. Reid and P. O. Vandervoort. Chicago University Press: Chicago.
Penrose, R. (1979). Singularities and time-asymmetry. In *General Relativity: An Einstein Centenary*, ed. S. W. Hawking and W. Israel. Cambridge University Press, Cambridge.
Penrose, R. (1981). Time-asymmetry and quantum gravity. In *Quantum Gravity 2*, ed. C. J. Isham, R. Penrose and D. W. Sciama. Oxford University Press: Oxford.
Penrose, R. (1983). Donaldson's moduli space: a 'model' for quantum gravity? In *Quantum Theory of Gravity*, ed. S. M. Christensen. Adam Hilger: Bristol.
Penrose, R. (1986). Gravity and state-vector reduction. In *Quantum Concepts in Space and Time*, ed. C. J. Isham and R. Penrose. Oxford University Press: Oxford.

Penrose, R. and Rindler, W. (1984). *Spinors and Space-Time*, Volume 1. Cambridge University Press: Cambridge.
Penrose, R. and Rindler, W. (1986). *Spinors and Space-Time*, Volume 2. Cambridge University Press: Cambridge.
Westfall, R. S. (1980). *Never at Rest*. Cambridge University Press: Cambridge.
Wheeler, J. A. (1964). In *Relativity, Groups and Topology: the 1963 Les Houches Lectures*, ed. B. S. De Witt and C. M. De Witt. Gordon and Breach: New York.
Wheeler, J. A. (1964). *Geometrodynamics and the Issue of the Final State in Relativite, Groupes et Topologie* (eds. C. deWitt & B. deWitt, Gordon and Breach, New York & London)
Wigner, E. P. (1961). Remarks on the mind-body question. In *The Scientist Speculates*, ed. I. J. Good. Heinemann: London.
Zel'dovich, Ya. B. (1966). *Sov. Phys. Dokl.*, 10, 771–2.

Note added in proof

It has been brought to my attention, by N. M. J. Woodhouse, that in Newton's manuscript fragment *De motu corporum in mediis regulariter cedentibus* (a precursor of *Principia* written in 1684) Newton had originally proposed to base his mechanics on five (or six) fundamental laws rather than the three that have come down to us through *Principia*. Law 4 was actually a clear statement of the (Galilean) relativity principle! Newton was well aware that these laws were not independent of one another, and for *Principia* he settled on the three that are now familiar to us. What is remarkable – apparently supporting the viewpoint tentatively put forward in this article – is that Newton had at one time, indeed, seriously contemplated using the relativity principle *as a fundamental principle* (despite what his views on 'absolute space' may or may not have been)!

175
ON QADIR'S INTERTWINING OPERATORS

Twistor Newsletter **25** (1987)

Asghar Qadir had been a Ph.D. student of mine when I was at Birkbeck College, London.

On Qadir's Intertwining Operators

In his London Ph.D thesis (~1970), Asghar Qadir suggested various twistor contour integral expressions, which enable one to pass from one representation of $gl(4,\mathbb{C})$ (or $su(2,2)$ or $SU(2,2)$ or $gl(n,\mathbb{C})$, etc.) to an equivalent one (given by different cohomology). These are closely related to expressions of Knapp & Stein, in relation to the Borel-Weil-Bott theorem. (I am grateful to R.J.B. and E.D. for making clear the connections.) Qadir's expressions deal with holomorphic functions $F(\mathbb{W},\mathbb{X},\mathbb{Y},\mathbb{Z})$ of four twistors, which are unaffected if any multiple of any of the later variables is added to any of the earlier variables, i.e.

(1): $F(\mathbb{W}+\lambda\mathbb{X}+\mu\mathbb{Y}+\nu\mathbb{Z}, \mathbb{X}+\rho\mathbb{Y}+\sigma\mathbb{Z}, \mathbb{Y}+\tau\mathbb{Z}, \mathbb{Z}) = F(\mathbb{W},\mathbb{X},\mathbb{Y},\mathbb{Z})$

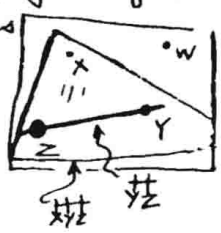
In differential terms, we have $(\forall\, \lambda,\mu,\nu,\rho,\sigma,\tau)$

$$\partial_{\mathbb{W}}F = \partial_{\mathbb{W}}F = \partial_{\mathbb{W}}F = \partial_{\mathbb{X}}F = \partial_{\mathbb{X}}F = \partial_{\mathbb{Y}}F = 0.$$

If we assume some homogeneity degree in each variable:

$$\partial_{\mathbb{W}}F = wF,\quad \partial_{\mathbb{X}}F = xF,\quad \partial_{\mathbb{Y}}F = yF,\quad \partial_{\mathbb{Z}} = zF,$$

then F is a cross-section of some line bundle over the flag manifold of projective twistor space, and could be rewritten as a function of $\mathbb{W}\mathbb{X}\mathbb{Y}\mathbb{Z}$ ($= \mathbb{IIII} \times \Delta$, where $\Delta = \mathbb{W}\mathbb{X}\mathbb{Y}\mathbb{Z}$), $\mathbb{X}\mathbb{Y}\mathbb{Z}$, $\mathbb{Y}\mathbb{Z}$ and $\mathbb{Z}$. (Special cases of such functions correspond to spinor fields on $\mathbb{CM}$ $\phi^{A\ldots L P'\ldots T'}_{\quad x\text{-}w\quad\; z\text{-}y}(x)$



$F = \lambda_A\lambda_B\cdots\lambda_L\pi_{P'}\cdots\pi_{T'}\phi^{AB\ldots LP'\ldots T'}(x)$, where λ_A, $\pi_{A'}$ are the projection parts of $\mathbb{WXYZ}$ ($\div\Delta$) and $\mathbb{Z}$, respectively, and where x^a corr. to $\mathbb{YZ}$.)

Qadir's expressions are (formal) contour integrals of the form

(A): $\oint F(\mathbb{W},\mathbb{X},\mathbb{Z},\mathbb{Y}+\lambda\mathbb{Z})d\lambda$, (B): $\oint F(\mathbb{W},\mathbb{Y},\mathbb{X}+\lambda\mathbb{Y},\mathbb{Z})$ (C): $\oint F(\mathbb{X},\mathbb{W}+\lambda\mathbb{X},\mathbb{Y},\mathbb{Z})\,d\lambda$

(D): $\oint F(\mathbb{W},\mathbb{Z},\mathbb{X}+\lambda\mathbb{Z},\mathbb{Y}+\mu\mathbb{Z})\,d\lambda\,d\mu$, (E): $\oint(\mathbb{Y},\mathbb{W}+\lambda\mathbb{Y},\mathbb{X}+\mu\mathbb{Y},\mathbb{Z})\,d\lambda\,d\mu$

(F): $\oint F(\mathbb{Z},\mathbb{W}+\lambda\mathbb{Z},\mathbb{X}+\mu\mathbb{Z},\mathbb{Y}+\nu\mathbb{Z})\,d\lambda\,d\mu\,d\nu$

In each case the resulting expressions satisfy (1). As for the homogeneity degrees, it is best to define the numbers of homogeneity $\omega = w$, $\xi = x+1$, $\eta = y+2$, $\zeta = z+3$.

We find that the expressions (A),(B),(C),(D),(E),(F) provide homogeneous functions with numbers of homogeneity given by, respectively, $(\omega,\xi,\zeta,\eta),(\omega,\eta,\xi,\zeta),(\xi,\omega,\eta,\zeta),(\omega,\eta,\zeta,\xi)$, $(\xi,\eta,\omega,\zeta),(\xi,\eta,\zeta,\omega)$, from which all permutations can clearly be built up. Qadir had previously explicitly worked out the Casimir operators, as explicit expressions in the symmetric functions of ω,η,ξ,ζ.

These contour integrals may be thought of as integrating representative functions for cohomology, and each contour circle reduces the cohomology by one step. Getting from a twistor function $f(z)$ to a massless field by the standard (generalized Bateman) contour integral arises as a special case:

$$\pi_{A'}\cdots\pi_{L'}\,\phi^{A'\cdots L'}(x) = \oint f(\dot{\gamma}+\lambda\dot{z})\,d\lambda$$

(this is the component in the "π-direction – which suffices, if we regard $\pi_{A'}$ as varying)

In accordance with twistor diagram procedures, etc. one should be able to reverse the reduction in cohomology by using open contours (i.e. contours with boundary). It would be worth while to explore this further, especially in order to relate to the following:

(I) The inverse twistor function (see R.P. in Quantum Gravity I; B.D.B., R.P. & G.A.J.S.). Here we have the standard procedure: twistor function → field, as an instance of (A) in a special case. F is independent of w, x and y, so we have $\partial_w F=\partial_x F=\partial_y F=0$. Applying (A) to F we get G, which now satisfies $\partial_w G=\partial_x G=0$, $\partial_z\partial_y G=0$. The inverse of this: field → twistor function, should be an open contour version of (A), i.e. (Ā), which looks awfully like the B.D.B. – R.P. – G.A.J.S. expression in Q.G. I.

(II) The twistor transform:

This should be an open (or open & closed, i.e. a cylinder) version of (F) ∴ homogeneity degrees are $-n,-1,-1,-1$ (if F has degrees $0,0,0,-n-2$, for helicity $\frac{1}{2}n$), i.e. $G=\oint\ldots$ is a function of $|\dot{x}\dot{y}\dot{z}|\div\Delta$ (= "$\mathcal{F}$"), of homogeneity degree $n-2$, multiplied by Δ^{-1} (as it should be).

Thanks due to A.P.H. and R.J.B.

References: R. Penrose (1974) Relativistic Symmetry Groups, in Group Theory in Non-Linear Problems, Ed A.O. Barut (Reidel)
A. Qadir, Ph.D thesis, Birkbeck College, London

176
ON THE ORIGINS OF TWISTOR THEORY

Gravitation and geometry: A volume in honour of Ivor Robinson (Monographs and Textbooks in Physical Science Lecture Notes **4**), 341–61. Bibliopolis Edizioni di Filosofia e Scienze, Naples, 1987

This article, in honour of Ivor Robinson's 60th birthday, provides some useful historical information concerning twistor theory in its earlier days.

672 On the origins of twistor theory. *Gravitation and geometry: A volume in honour of Ivor Robinson* (Monographs and Textbooks in Physical Science Lecture Notes **4**), 341–61, 1987

ON THE ORIGINS OF TWISTOR THEORY

ROGER PENROSE
Mathematical Institute
Oxford, U.K.

1. Introduction

It is an especial pleasure for me to have this opportunity to pay my respects to my friend and colleague Ivor Robinson. I have chosen to hold forth on the origins of twistor theory for two reasons. The first is that Ivor's 60th birthday very nearly coincides with what is (for me) the 20th birthday of the theory, so this seemed to be a reasonable point at which to examine how it stands today, in relation to its original aims and aspirations. But also Ivor himself played a direct and important part in those origins. In fact, I can think of several essentially independent influences that Ivor had, one of which was quite crucial. I hope the reader will bear with me and forgive me for presenting an account which, to a large extent, consists of personal or technical reminiscences. But I hope, also, that there will actually be some scientific value in these ramblings.

I should first make clear what I mean, here, by the "origins" of twistor theory. I am referring to the origins of my own rather specific approach to a physical theory. I appreciate that many of the ideas go back very much farther than twenty years. Most particularly, Felix Klein put forth his correspondence between the lines in complex projective 3-space and a general quadric in projective 5-space as long ago as 1870 (Klein 1870, 1926), this correspondence being based on the coordinates of Julius Plücker (1865, 1868/9) and Arthur Cayley (1860, 1869), (or of Hermann Grassmann, even earlier). Sophus Lie had noted essentially the key "twistor" geometric fact that oriented spheres in complex Euclidean 3-space (including various degenerate cases) could be represented as lines in complex projective 3-space (contact between spheres represented as meeting of lines) already in 1869 (*cf.* Lie &

W. Rindler and A. Trautman (Ed.)
GRAVITATION and GEOMETRY
a volume in honour of I. Robinson
BIBLIOPOLIS, NAPLES 1987

Scheffers 1896) as was pointed out to me by Helmuth Urbantke some years ago. The spheres may be thought of as the $t = 0$ representation of the light cones of events in Minkowski space, so the Lie correspondence *in effect* represents the points of (complexified compactified) Minkowski space by lines in complex projective 3-space, where meeting lines describe null-separated Minkowski points — the twistor correspondence! The local isomorphism between the "twistor group" SU (2,2) and the connected component of the group O (2,4) was explicitly part of Élie Cartan's (1914) general study and classification of Lie groups. The physical relevance of O (2,4) in relation to the conformal motions of (compactified) Minkowski space-time had been exploited by Paul Dirac (1936 b) and the objects which I call twistors (namely the spinors for O (2,4)) had been explicitly studied by Murai (1953, 1954, 1958) and by Hepner (1962). (See also Gindikin 1983 for a discussion of these matters.)

Moreover, as Ivor Robinson pointed out to me some ten or so years ago, a certain line-integral expression for representing the general (analytic) solution of the wave equation in terms of holomorphic functions of three complex variables was known to Bateman in 1904 (see Bateman 1904 and 1944, p. 96), this having arisen from a similar expression due to Whittaker (1903) for solving the three-dimensional Laplace equation, and Bateman also gave a similar line-integral expression for solving the free Maxwell equations (Bateman 1944, p. 100). By a simple transformation of variables, these become the helicities zero and one cases of the basic contour integral formula (Penrose 1968, 1969 a) giving the linear field case of the so-called "Penrose transform" of twistor theory. The Radon transform (Radon 1917, Gel'fand Graev & Vilenkin 1966) and its generalizations may also, from a different angle, be regarded as providing models for (and generalizations of) this twistor expression. In addition, the classic Weierstrass (1866) construction (cf. Darboux 1914) (*which was known to me!*) *had provided a paradigm for the explicit* solution, in terms of free holomorphic data, of an important non-linear problem (Plateau's problem). This may be regarded as a direct antecedant of the later non-linear twistor constructions for (anti-) self-dual gravitational (Penrose 1976; cf. also Hitchin 1979) and Yang-Mills fields (Ward 1977, Atiyah & Ward 1977, Aityah, Hitchin, Drinfeld & Manin 1978).

Much of this previously existing material was not known to me

On the origins of twistor theory. *Gravitation and geometry: A volume in honour of Ivor Robinson* (Monographs and Textbooks in Physical Science Lecture Notes **4**), 341–61, 1987

twenty years ago. But the Klein correspondence was something I had been well acquainted with since my undergraduate days. So some might argue that there was not a great deal left to be original about in the basic twistor scheme. Nevertheless I do feel that I have a good claim to some sort of originality! This — if we discount a fair number of (non-trivial) later mathematical developments — lies primarily in the essential "physical idea" that the *actual* space-time we inhabit might be significantly regarded as a *secondary* structure arising from a deeper twistor-holomorphic reality. The basic idea, it could be argued, provides little more than a shift in viewpoint, but it is this shift that provides crucial motivation and it also gives, in a sense, whatever *physical* content the theory has had, so far. This viewpoint has guided us in certain unexpected and often fruitful directions, providing some surprising mathematical insights and descriptions of basic physical fields and concepts. It has enabled us to achieve results that had not seemed possible to achieve by more conventional procedures. (For accounts of some of these, see Hodges *et al.* 1980). Nevertheless, twistors do *not*, as yet, provide a new physical theory in the usual sense that predictions different from those given by conventional procedures are yet forthcoming. However, in order that the above "physical idea" should have genuine *physical* content, it must at *some* stage lead to a successful physical theory in this sense — or else be consigned to the dustbin ([1]) of scientific history!

2. Some background ideas

Let me try to set in perspective my own state of mind some twenty years ago, and to explain some of the reasons why I felt that a different viewpoint with regard to space-time structure, of the kind provided by twistor theory, was needed. I had, for a good many years earlier, been of the opinion that the space-time continuum picture of reality would prove inadequate on some small scale. I do not propose to discuss all the reasons for this — and in any case it is a view that is hardly original with me. Indeed, that the quantum nature of reality should affect the very structure of space-time at some scale is now a more-or-less accepted viewpoint among those physicists who have examined this question in some depth (cf. Schrödinger 1952, Wheeler 1962). But I think that

most physicists would believe that such effects should be relevant only at the absurdly small quantum gravity scale of 10^{-33} cm (or smaller). My own attitude was somewhat different from this. While it might be that only at 10^{-33} cm is it *necessary* to invoke a description of space-time radically removed from that of a manifold, my view was (and still is) that even at the much larger levels of elementary particles, or perhaps atoms, where quantum behaviour holds sway, the standard space-time descriptions have ceased to be the most physically appropriate ones, and some other picture of reality, though at that level equivalent to the space-time one, should prove to be the more fruitful. The very fact that quantum behaviour *is* so hard to picture in the normal way had seemed to me to argue stongly that the normal space-time picture of things, even at that level, is inappropriate physically. Indeed, there was nothing really new in this either, inadequacy of space-time *pictures* being very much a part of the standard quantum-mechanical philosophy. However, I felt that one ought to try to be more positive than this, in actually providing a picture of objective reality, albeit one perhaps radically different from the usual one.

Space-time descriptions of the normal kind can, of course, be used at the atomic or particle level provided that the quantum rules are correctly applied, and they have implications that are extraordinarily accurate. Thus, this new geometrical picture must, at that level, be mathematically *equivalent* to the normal space-time picture — in the sense that some kind of mathematical transformation must exist between the two pictures. However, the new description ought to incorporate quantum behaviour more readily and naturally than the old. Moreover, at the quantum gravity level of 10^{-33} cm, or at the level of space-time singularities, it ought to provide an essentially different and more accurate picture of things.

3. Motivation from spin-network theory

These considerations were among those that motivated the twistor approach, but the extent to which these particular aims have been actually fulfilled remains somewhat problematical as of now. Another initial motivation whose present status is unclear is that from spin-network theory (originating in about 1958 and published later: Penrose 1969 b, 1972 b; cf. Moussouris 1983). I had had the idea, from the

time I was a Research Fellow at Cambridge in the late 1950's, that the discrete combinatorial rules obeyed by quantum-mechanical (total) angular momentum could be used as a starting point for building up space-time structure. Though grounded in accepted quantum-mechanical principles, spin-network theory provided a picture of space which is entirely combinatorial in nature, so long as only a finite number of particles (or "units") are involved. Only in the limit, when this number becomes infinite, did a continuous picture of space arise. (These ideas were stimulated to a large extent by long discussions in Cambridge with Dennis Sciama about Mach's principle: "What happens to the concepts of space and direction if all the matter in the universe is removed save a small finite number of particles?") The spin-network scheme worked well enough (and in some respects surprisingly well) for a non-relativistic scheme in which directional concepts — though not spatial displacements — arise. The scheme was based on the representation theory of SO (3) (or, more correctly, of SU (2)), and for a more complete relativistic scheme, which included also spatial displacements, it had seemed that the (restricted) Poincaré group should replace SO (3) in the discussion. I had been unable to overcome certain obstacles to this (my former student John Moussouris having only quite recently been able to deal promisingly with them, *cf.* Moussouris 1983) and I had largely set the question of a direct generalization of spin-network theory aside.

However, in a slightly indirect way, spin-network theory did have a conceptual influence on the development of twistors. Moreover, there is perhaps a little irony in a certain "false start" which occurred in the winter of 1959/60, while I was at Princeton. I had felt that, in the context of the development of a suitable relativistic spin-network theory, it might prove fruitful to examine the total momentum-angular momentum structure of a relativistic physical system in the cases where the 4-momentum was a *null* vector, the view being that this was more primitive than the standard timelike case. It seems that I had effectively worked out what was needed, and though the resulting objects are actually essentially twistors, the proper realizations had not then come to me and I did not develop the ideas further at the time. Only towards about 1970, long after twistors had been subsequently developed by quite another route, did I realize that they had also this null momentum-angular momentum interpretation (*cf.* Penrose & MacCallum 1972).

As things have worked out so far, twistor theory has not moved

much in the "combinatorial" direction of spin-networks. Instead, the (seemingly) very different complex-analytic aspects of twistors have been the ones that have proved to have greatest importance. The one place where the possibility of a connection with spin-network theory remains fairly strong is in twistor diagram theory (Penrose & MacCallum 1972, Penrose 1975 a, Sparling 1975, Hodges & Huggett 1980, Hodges 1983, 1984) and in a certain sense it has been SU (2,2) rather than the Poincaré group which has so far replaced SO (3) in the discussion. Much more work will be needed to discern whether these relationships are more than superficial.

4. Holomorphicity in classical space-time structure

Other motivations than these played much more direct roles in the development of twistor theory. Spin-networks notwithstanding, the role of *complex numbers* in quantum theory had long struck me as a quite crucial one. If the "correct" geometry for the world is to be a closely quantum one, then these same complex numbers must be an essential part of this geometry. My training as a (largely pure) mathematician had taught me something of the power, subtlety and elegance of complex (holomorphic) geometry. It had seemed fitting that this might be the geometry most basic to the structure of the physical world. Yet in its most obvious manifestations, physical geometry *seems* to be geometry over **R**, not **C**.

Nevertheless certain hints of a complex underlying structure had been apparent to me for some time. The fact that the null directions at a point have the holomorphic structure of a Riemann sphere (*cf.* Penrose 1959) had long impressed me, this fact being closely related to the complex nature of Lorentzian spinors. It had, for a long time, seemed to me that spinors, and particularly Lorentzian 2-spinors, are more fundamental than Minkowskian world-vectors, and that the latter should be regarded as derived from the former ([2]). (Compare Rzewuski 1958).

In addition to this, complex numbers often have significant roles to play in solutions of Einstein's vacuum equations. I had been particularly impressed by the nature of the plane-fronted waves that Ivor had introduced me to (these being due originally to Brinkmann 1923, but later rediscovered by Ivor Robinson, *cf.* Robinson 1956) and also of their later generalizations to spherically fronted waves — due to Robinson

and Trautman (1962). For these waves the behaviour along the null geodesics of propagation is fixed and there is completely arbitrary variation from wave-front to wave-front. But within each wave-front the strength and polarization of the wave is governed by a single arbitrary holomorphic function. What had struck me was the direct appearance of a free *holomorphic* function in the solution, the modulus and argument of this function *both* playing a direct role (as strength and polarization, respectively). This kind of feature had led me to believe in some role for holomorphic structure lying "behind the scenes" in solutions of Einstein's (vacuum) equations.

5. The positive-frequency condition

All this had concerned a possible role for "holomorphic structure" governing space-time even at the classical level. But I was searching for something in which *quantum* mechanical ideas arose in a unified way in relation to space-time geometry. It had been Ivor's (and my) good friend Engelbert Schücking who had impressed upon me, at Syracuse N.Y. in the spring of 1961, the fundamental importance of the positive-frequency condition in complex solutions of field equations, as the hallmark of quantum field theory. He had also persuaded me of the significance of conformal invariance, in relation, most particularly, to the massless fields of each spin (*cf.* Cunningham 1910, Bateman 1910, McLennan 1956; fields defined by Dirac 1936 a). In my attempt to come to terms with the positive-frequency concept, especially in the context of general relativity where Fourier analysis does not find a comfortable home, I had been led to appreciate the value of formulations where this concept is expressed in terms of analytic continuation into *complex* regions of space-time. An added bonus was that this type of formulation can readily be manifestly conformally invariant, whereas the momentum-space approach afforded by Fourier analysis is particularly obscure in this respect.

6. Massless fields

I had been struck by what had tantalizingly seemed to be possibly a similarity (or relationship) between the analytic continuation properties

needed for the positive-frequency condition and a certain generalization, to massless fields of arbitrary spin, of the Kirchhoff-D'Adhemar integral expression for solving the wave equation. I had found this generalization a short time earlier, probably while at Princeton in 1960 (*cf.* Penrose 1980, Newman & Penrose 1968, Penrose & Rindler 1984) and its conformal invariance properties had, I think, become apparent to me before the autumn of 1963.

I think that I had very much come around to the view that *massless* particles and fields were to be regarded as more fundamental than massive ones. My own reasons for believing this were probably very much bound up with the idea that 2-spinors should be regarded as more fundamental than Minkowski world-vectors, since it is a *null* vector (the 4-momentum of a *massless* particle), rather than a timelike one, which arises naturally from a single spin-vector of the 2-spinor formalism (*cf.* Penrose & Rindler 1984). Moreover, I had been particularly struck by the elegant properties of the massless free fields of arbitrary spin (in Minkowski space **M**), the conformal invariance properties that Engelbert had drawn my attention to, having contributed considerably to my appreciation of this elegance.

It had also seemed to me that massive particles and fields should probably be, in some sense, built up from massless ones. The van der Waerden form of the Dirac equation is rather suggestive of the mass playing a role as a coupling between two Dirac-Weyl neutrino-type fields (*cf.* Penrose 1968 — but this idea is undoubtedly not original with me). In the summer of 1961 I had generalized my Kirchhoff-D'Adhemar type expression to handle the Dirac equation looked at in this way, and in 1962, the Maxwell-Dirac equations. The formulae were all sums of integrals over spaces of zig-zag and forked collections of *null* straight segments (Penrose & Rindler 1984). Thus, in this approach, *all* fields, whether massless or massive, are viewed as propagating along light cones, and I liked to take the view that, in some sense, only the *null* directions were really "there"!

In accordance with this, it had seemed to me that it was important to search for some kind of background formalism, in which massless fields — subject to the positive-frequency condition of quantum field theory — would play some sort of primitive role.

Moreover, the geometric framework of such formalism ought to be understood.

7. Compactified Minkowski space (complexified)

I had already thoroughly explored the structure of conformally compactified Minkowski space in the spring of 1962 (*cf.* Penrose 1965, the much earlier work of Bôcher 1914, Coxeter 1936, Kuiper 1949, Rudberg 1958 and, no doubt, Lie, Möbius, Cartan and others, having been, in detail, unknown to me at the time, though certain crucial ideas of "inversive geometry" had filtered through to me) and I had become convinced that some form of unified (but convoluted?) complex viewpoint should exist, whereby the analytic extensions needed for the positive-frequency condition, the generalized Kirchhoff-D'Adhemar formula, null lines, spinors and the holomorphic structure of the space of null directions, the appearance of holomorphic functions in solutions for (linearized) gravity, the conformal invariance properties and the compactification, would all be among its manifestations. Apparently Engelbert Schücking had himself been thinking along somewhat related lines since he had repeatedly hinted to me, in the autumn of 1963 in Austin Texas (where he, I, Roy Kerr, Ray Sachs, Jerry Kristian and others had joined Alfred Schild's Relativity Center) that some kind of alternative "complex view" of space-time might prove fruitful.

Of course the possibility of simply describing things in terms of *complexified* (compactified) Minkowski space **CM** had occurred to me but — for reasons which are still not entirely clear to me — I had (correctly) rejected this as insufficiently subtle for Nature. I think that one reason for being unhappy with **CM** as playing a *primary* role in physics was that the complexification is far too gross. As many additional "unseen" dimension (namely four) would need to be adjoined as are already directly physically interpretable (^{3}).

It had seemed to me that one needed to find some higher-dimensional analogue of the way that the real axis divides the complex plane into two halves, functions holomorphic in one half providing the positive frequency functions on **R** and those holomorphic on the other half, the negative frequency functions. Minkowski space **M** forms a tiny 4-real-dimensional subspace of the 8-real-dimensional manifold **CM** and therefore is, by itself, incapable of dividing **CM** into two halves. I had been aware of the two "tube" domains $\mathbf{CM}^+$ (the forward tube — given by position vectors whose imaginary parts are past-timelike) and $\mathbf{CM}^-$ (the backward tube — imaginary parts future-timelike) into which the

positive frequency and negative frequency fields, respectively, extend holomorphically. But **M** is not a boundary of these domains (in the ordinary sense) and there is also a large open region in **CM** (imaginary part spacelike) lying "between" $\mathbf{CM}^+$ and $\mathbf{CM}^-$. Fortunately I was not aware, at the time, of the concept of *Shilov boundary* — the sense in which **M** *can* be regarded as a kind of boundary of each of $\mathbf{CM}^+$ and $\mathbf{CM}^-$ — for if I had been, I might perhaps have been tempted, after all, in that kind of direction!

So I had this feeling that some sort of complex space was needed, which somehow fell naturally into two halves, the common boundary between these two halves being, like the real axis of the complex plane, the more readily physically identifiable "real" part of the space. In this way, a good part of the holomorphic structure that seemed already to be "visible" in relativistic physics might be interpretable directly in terms of the induced partial complex structure (or CR-structure, as I learned later) on this common boundary. I had no idea what this complex space was, nor, indeed, whether or not its very existence might be merely a product of wild fantasy.

8. Robinson congruences and twistors

It was in such frame of mind that I recall being driven ([4]) from San Antonio to Austin Texas (by Pista Oszváth) following a weekend family outing with Rindlers and Oszváths, the others following in a later car. I began thinking about an ingenious construction that Ivor had, some while earlier, described to me. He had been endeavouring to find null (real) solutions of Maxwell's free-space equations which were everywhere non-singular and of finite total energy ([5]). His procedure was to base such a solution on a certain twisting shear-free congruence of *rays* (null straight lines) in **M** — now referred to as a *Robinson congruence*. To construct Ivor's congruence consider first the family (congruence) of rays x meeting a given ray q. This is a *special* Robinson congruence. The tangent directions to the various rays x at the various points of each x provide a shear-free geodetic field of null directions which is, however, singular along q. Next perform an arbitrary complex translation, or indeed any complex Poincaré transformation, to this configuration (complexified). At each point of **M** we now have a *complex* null direction. But a complex null direction determines two real null directions. In

spinor terms (*cf.* Penrose & Rindler 1984), if the complex null direction is that of $\alpha^A\beta^{A'}$, these two real null directions are those of $\alpha^A\bar{\alpha}^{A'}$ and $\beta^A\beta^{A'}$. Select (say) the former, consistently. We thus have another real field of null directions on **M** which again turns out to be geodetic and shear-free. But now this field is (in general) everywhere non-singular (and twisting). The family of rays whose tangent directions constitute this field is a Robinson congruence.

It is reasonably clear from the construction that Robinson congruences constitute a holomorphic family (i.e. the manifold whose points are the different Robinson congruences is naturally a complex manifold) essentially because they arise from general holomorphic motions of **CM**. Among these congruences are the special Robinson congruences, each of which may be viewed as a way of describing a particular ray q. The general Robinson congruences describe, in a sense, "complexified" rays.

I had, somewhat earlier, worked out the geometry of a general Robinson congruence: in each time-slice t = const. of **M** the projections of the null directions into the slice are the tangents to a twisting family of linked circles (stereographically projected Clifford parallels on S^4 — a picture with which I was well familiar), and the configuration moves with the speed of light in the (negative) direction of the one straight line among the circles. (See fig. 1).

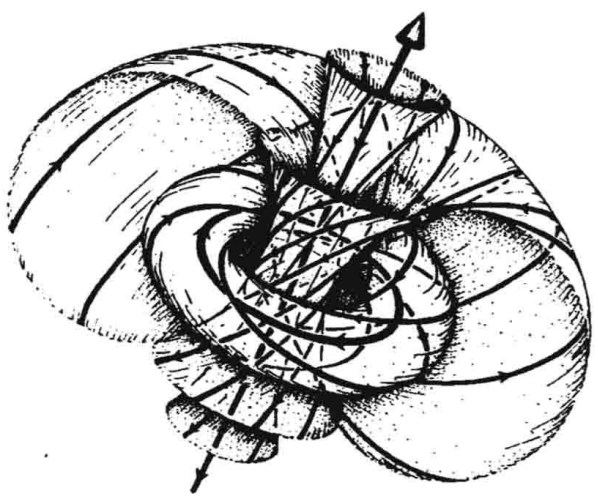


FIGURE 1: A time-slice (t = 0) of a Robinson congruence.

I decided that the time had come to count the number of dimensions of the space $\mathcal{R}$ of Robinson congruences. I was surprised to find, by examining the freedom involved in fig. 1, that the number of real dimensions was only *six* (so of only *three* complex dimensions) whereas the special Robinson congruences, being determined by single rays, had *five* real dimensions. The general Robinson congruences must twist either right-handedly or left-handedly, so $\mathcal{R}$ had two disconnected components $\mathcal{R}^+$ and $\mathcal{R}^-$, these having a common five-dimensional boundary $\mathcal{S}$ representing the special Robinson congruences. The complex 3-space of Robinson congruences was indeed divided into two halves $\mathcal{R}^+$ and $\mathcal{R}^-$, by $\mathcal{S}$.

I had found my space! The points of $\mathcal{S}$ indeed had a very direct and satisfyingly relevant physical interpretation as "rays", i.e. as the classical paths of massless particles. And the "complexification" of these rays led, as I had decided that I required, to the adding merely of one extra real dimension to $\mathcal{S}$, yielding the complex 3-manifold $\mathbf{PT} = \mathcal{S} \cup \mathcal{R}^- \cup \mathcal{R}^+$.

After returning home (in Austin) I was able to employ the spinor techniques that had long been familar to me and it did not take me long to realize that **PT** was indeed a complex projective 3-space ($\mathbf{CP}^3$) the lines within which having **CM** as their Klein representation. This was perhaps a somewhat roundabout route to the picture of **CM** as a Klein quadric. I had been well aware for some time that **CM** was indeed a non-singular complex 4-quadric, and for considerably longer that non-singular complex 4-quadrics are always Klein representations of lines in some $\mathbf{CP}^3$. But without such a roundabout route, would the "absurd" thought of combining these two ideas ever have occurred to me before? (Curiously, the answer to this question is "yes". Apparently I had done this about nine months earlier, but the odd-looking reality conditions on the Cayley-Plücker coordinates and lack of physical motivation at that time had put me off. So I completely forgot about it until going though my notebooks again not many days ago!)

The various physical motivations that I had collected together were, indeed, crucial for me, in order that the necessary ingredients of this strange idea should come to me and, more importantly, take hold.

 On the origins of twistor theory. *Gravitation and geometry: A volume in honour of Ivor Robinson* (Monographs and Textbooks in Physical Science Lecture Notes **4**), 341–61, 1987

9. The Kerr theorem

It is one of those odd coincidences, which seem to happen more frequently than they should, that a day or so later I overheard Roy Kerr (who had an office on one side of mine, Engelbert Schücking having had the office on the other side) explaining his method of obtaining *all* (analytic) shear-free geodetic null congruences in **M** to Ray Sachs. I had him explain it to me and was amazed to find that the coordinates that came up naturally in his construction were precisely those "twistor coordinates" which I had just found as the projective coordinates for **PT** (projective twistor space). So almost at once a new role for my complex space was at hand: the general (real-analytic) shear-free ray congruence in **M** is provided by the intersection of **PN** ($= \mathcal{S}$, the space of rays in **M**) with a general complex-analytic (holomorphic) 2-surface $\mathcal{K}$ in **PT** (see Penrose 1967).

Indeed, this had further ramifications. It was not long before I realized that another of Ivor's results had an immediate interpretation in terms of **PT**. He had shown that the general (real-analytic) *null* solution of *Maxwell's equations* could be described in terms of a general (analytic) shear-free geodetic congruence, with just the additional information of one free holomorphic function F, of two complex variables. In effect, the variables labelled the rays of the congruence, and in **PT** the function F turned out to be holomorphic on $\mathcal{K}$. So one was presented with a very neat "holomorphic" description of null Maxwell fields in **PT**.

10. General massless fields in **M**

This elegant picture generalized in an obvious way to null massless fields of higher (integral or half-integral) spin and, in particular, as Ivor noted, to linearized Einstein fields. So at least in the linear limit, my motivation from the "holomorphic" structure of plane-fronted and spherically fronted waves (both of which are indeed null) was satisfied. But I needed more than this. My motivations required than *general* solutions of the massless free field equations be simply representable in terms of twistors.

I tried many formulations, but these, being effectively only straightforward transcriptions of the space-time fields and equations, lacked sufficient elegance and naturality. Functions of Cayley-Plücker coordin-

ates presented some awkwardness because of the identical relations they satisfy — though I found a reasonable enough formulation in these terms. The alternative seemed to be two-point functions in **PT** and this, also, seemed not quite what was needed. It had occurred to me that the generalized Kirchhoff-D'Adhemar expression somewhat resembled a Cauchy integral formula, and I had it in mind that, in some sense, the massless free field equations ought to be the Cauchy-Riemann equations in disguise. "Would it not be most fitting" I had thought to myself "if a massless field could be described by a single *free* holomorphic function f on **PT**?" The freedom in the solutions that would occur should be just right. **PT** (for *holomorphic* information) counts as three-dimensional, which is the same dimensionality as an initial data surface in **M**, from which fields would uniquely propagate. The hope, then, was that the field equations should simply evaporate!

Perhaps this was just a pipe-dream — but in due course the thought occurred to me (not until 1966, in London, though!) that, in the null case, Roy Kerr's surface $\mathcal{K}$ in **PT** should be treated as a *pole* of the putative "f", where Ivor's holomorphic function F would be its residue, the field being obtained by contour integration of f. (It was the presence of certain singularities in the field, and the fact that these would duly result from pinching of such contours — owing to a tangency with $\mathcal{K}$ — which suggested this.) The formula (Penrose 1968, 1969 a) then dropped out. It was many years later that Ivor (and Lane Hughston) pointed out to me that Bateman had, in effect, found essentially the same formula over sixty years earlier! My own formulation did, however, have the advantage that geometric and conformal transformation properties of the field were made manifest, and also connections with the (above) Robinson-Kerr theorem, etc., were exhibited.

11. Cohomology

But what about the positive-frequency condition? I had been motivated by the hope that positive frequency fields were to be associated with holomorphic extension into $\mathbf{PT}^+$ ($= \mathcal{R}^+$, the space of right-handed Robinson congruences), and negative frequency fields, extension into $\mathbf{PT}^-$ ($= \mathcal{R}^-$, the space of the left-handed Robinson congruences). The condition turned out, instead, to be that f had *separated singularity*

sets in $\mathbf{PT}^+$, or $\mathbf{PT}^-$, according as the resulting massless field was to be of positive or negative frequency. What had happened to my "crucial" motivation that **PT** needed to be naturally divided into two halves by **PN** ($= \mathcal{S}$)?

The answer had to wait ten years (until the spring of 1976), after I had acquired some appropriate re-education from Michael Atiyah! It turned out that f was to be interpreted as a representative cocycle for an element of a holomorphic (first) *sheaf cohomology group*. I call such an element a holomorphic 1-function. Then, the 1-function does indeed extend globally to $\mathbf{PT}^+$ or $\mathbf{PT}^-$ according as the resulting massless field is of positive or negative frequency.

I have never been able to decide whether I was actually "insightful" here or just "lucky"! I had come to feel that there perhaps was something a little phoney about my "crucial" motivation, and had realized that neither $\mathbf{PT}^+$ nor $\mathbf{PT}^-$ could globally admit non-trivial *ordinary* holomorphic functions (0-functions). I was unaware (or had forgotten what little I knew) of holomorphic sheaf cohomology and had learned to live with the concept of "separated singularities". But then, after thirteen years this motivation was finally satisfied after all, in a quite unexpected way!

12. Space-time curvature

I suppose that it was the many attractive properties of twistor geometry which mainly held me to that picture for so long. Even so, I recall having had, back in 1964, some considerable difficulty in persuading my colleagues, (Engelberg Schücking apart) of the significance or possible fruitfulness of this odd viewpoint, regarding space-time, that I was attempting to put forward. I expect that part of the reason for this was that my colleagues were basically all general-relativists (though, no doubt, I should have perceived even *less* enthusiasm had they been particle physicists or quantum field theorists!), my ideas seeming to be completely tied to the building up of *flat* (or at least conformally flat) space-time. Indeed, looking back on all this I find it difficult to see why I was not myself more disturbed by this lack of curvature than I seem to have been. For a good many years earlier I had been (and remain to this day) firmly convinced of the (conformal) curvature of

physical space-time. I think that I had been able to keep my worries at bay, at first, by the thought that such twistor ideas referred only to the quantum level of things, and that at this level space-time could be treated as flat, for the most part, the curvature perhaps arising only in discrete "quantized" lumps. I had been somewhat influenced by some of Tullio Regge's ideas for building effectively curved spaces in polyhedral fashion out of flat pieces (Regge 1961, Penrose 1972 a) and had imagined that twistor ideas might be carried over to such situations in a piecewise manner. Somewhat later, while at Cornell in 1967, I was actually able to follow up this kind of idea in a concrete way. By examining plane-fronted, and later spherically fronted, gravitational waves with δ-function curvature — so the space-time was flat on either side of the wave — it proved to be possible to provide a twistor description involving a "shift" in the structure of the twistor space (Penrose 1968, Penrose & MacCallum 1972; *cf.* Bondi, Pirani & Robinson 1959). The very holomorphic functions (referred to above) which describe the structure of the curvature along the wave-front, proved to have significance here and had valuable implications in later developments in twistor theory.

In more recent years I have become less happy about the above point of view whereby space-time curvature arises in some sort of "quantized" discrete bits. Fortunately some newer developments, not unrelated to those described above, have arisen. Rather than "piecing together" the *space-time* from separate flat pieces, it is the *twistor space* which must, in an appropriate sense, be so pieced together. The local holomorphic structure of twistor space allows just the right sort of flexibility to enable such piecings to encode, in the *global* structure of the resulting twistor space, the appropriate (local) curvature information for the resulting space-time (Penrose 1976, Penrose & Ward 1980; an important input had been the $\mathcal{H}$-space ideas of Newman 1976). This curious encoding of local space-time information in global twistor structure has emerged as a characteristic feature of the representation of physical fields in twistor terms. Indeed, this is already a feature of the holomorphic 1-function description of massless linear fields in **CM**. And the transition functions for the above piecing together of twistor-space portions provide, in effect, a kind of non-linear 1-function.

 On the origins of twistor theory. *Gravitation and geometry: A volume in honour of Ivor Robinson* (Monographs and Textbooks in Physical Science Lecture Notes **4**), 341–61, 1987

13. Self-dual and anti-self-dual fields

In fact the above type of construction has been successful in producing the general anti-self-dual solution of Einstein's vacuum equations (Penrose 1976). A corresponding subsequent construction due to Richard Ward (1977) provided the general anti-self-dual solution of the Yang-Mills equations. More accurately, these are non-linear version not of my original 1-functions, but of a slight modification suggested (in effect in 1973) by Lane Hughston. This produced anti-self-dual linear fields (with standard twistor conventions) whereas my original formulation produced self-dual fields. With the positive frequency condition these are the negative- and positive-helicity fields, respectively (*cf.* Penrose & MacCallum 1972).

For many years now I have been struggling to find a non-linear version of my *original* 1-function description in the gravitational case (the "googly graviton" — which, as cricketers know, spins right-handedly though the bowling action suggests the reverse). I believe that it is likely a full solution is close — though there remains the possibility that all will collapse. The hope is that combining these "googly" ideas with the earlier ("leg-break") ones will eventually result in a complete formulation of the Einstein equations. Indeed, I feel that such grandiose objectives must necessarily be achieved if twistor theory is ultimately to realise its aims whereby space-time concepts can all be supplanted by twistor-holomorphic ones. We live, as before, in hope.

The idea that massless fields can usefully be split into their self-dual and anti-self-dual parts is clearly a crucial one in all this. Where first did I hear of this concept? I think that it was on the steps of some building in Amsterdam, at the International Congress of Mathematicians in 1954. Someone was explaining to me, with characteristic enthusiasm and flair, what had seemed to me at the time to be a very odd idea indeed: that one should add the Maxwell field tensor to i times its dual! That someone was Ivor Robinson (6).

14. Other developments

It is hard for me to assess the success, or lack of it, that twistor theory has had to date. Much has happened in twenty years, though it has seemed for most of the time that progress has been grindingly slow.

The mathematics has turned out to be much more difficult than I had guessed, and many of the necessary ideas, more sophisticated. That is my excuse for the slowness. But the programme has not as a whole been turned back, and that is encouraging. Some parts of the programme, though initially promising, have not moved a great deal in recent years, perhaps because of a lack of some essential new insight. (The particle programme, *cf.* Penrose 1975 b, Hughston 1979, Perjés 1975, 1979, Perjés & Sparling 1979 seems to be one of these.) Others have begun to move again only very recently — and here I refer particularly to the theory of twistor diagrams, into which some very promising-looking new ideas have been injected by Andrew Hodges (1985) within the last year or so. It will be exciting to see where these lead.

There have also been some unexpectedly encouraging developments which I might refer to as "spin-off" rather than part of what I had thought of as the main objectives. I would include the solution of the Euclidean Yang-Mills "instanton" problem (Atiyah, Hitchin, Drinfeld & Manin 1978) and the construction of positive definite Ricci-flat self-dual 4-manifolds (*cf.* Hitchin 1979) among these; also "Killing spinors" (*cf.* Walker & Penrose 1970), non-realizable CR-structures (*cf.* Penrose 1982 a) and my own construction of a "quasi-local mass and angular momentum" complex, which provides a twistor-motivated definition of these quantities as surrounded by general spacelike topological 2-spheres in arbitrary space-times (Penrose 1982 b *cf.* also Tod 1983, Shaw 1983, Penrose & Rindler 1985). The status of this definition is not yet completely clear, but at worst it seems to be a considerable improvement on all earlier suggestions, and some results provide good cause to be optimistic.

Whatever the future provides for twistor theory, it should be interesting.

([1]) Perhaps this seems unduly harsh. An idea — Hamiltonian theory, for example — may have immense utility and lead to new insights without, in this sense, having any new physical content. Thus in *this* sense it was, I suppose, the advent of quantum theory which saved the Hamiltonian viewpoint from the dustbin!

([2]) It is not appropriate for me to dwell here at length on all the numerous other motivations, some vague and some fairly clear-cut, which influenced the direction of the development of twistors. Among these was a desire for a formalism *tailored* to the four-dimensional $(+ - - -)$ structure of our space-time, rather than for something not so specific. The holomorphic nature of the space of null directions, in our particular dimension and signature, seemed to be a highly suggestive clue. Other motivations were provided by the experimental facts of left-right asymmetry and non-locality (*cf.* Lee & Yang 1956; Bohm 1951, Aharonov & Bohm 1959). Twistors have emerged as very compatible with these objectives.

([3]) Perhaps in the present climate of eleven-dimensional generalized Kaluza-Klein theories this objection would carry little weight with most people. However, to me it was, and still is, a fundamental drawback.

([4]) Apparently on 1, December 1963 - for which date I thank Zsuzsi Ozsvath.

([5]) He had not quite succeeded in satisfying this final requirement, but a slight modification of his solution (found a few months later by twistor-type methods, *cf.* Penrose 1965, provided what he had been seeking. We now refer to such solutions as *elementary states* (see Penrose 1975 a) and they have importance in twistor theory.

([6]) He was probably also explaining to me about self-dual *null* bivectors, but their relationship to spinors seems to have been something I learnt later!

REFERENCES

Aharonov, Y. & Bohm, D. (1959), Phys. Rev. **115**, 485.

Atiyah, M. F., Hitchin, N. J., Drinfeld, V. G. & Manin, Yu. I. (1978), Phys. Lett., **65A**, 185.

Atiyah, M. F. & Ward, R. S. (1977), Commun. Math. Phys., **55**, 111.

Bateman, H. (1904), Proc. Lond. Math. Soc. (2) **1**, 451.

Bateman, H. (1910), Proc. Lond. Math. Soc., (2) **8**, 223.

Bateman, H. (1944), Partial Differential Equations of Mathematical Physics (Dover, New York).

Bôcher, M. (1914), Bull. Amer. Math. Soc., **20**, 185.

Bohm, D. (1951), Quantum Theory (Prentice-Hall, Englewood Cliffs).

Bondi, H., Pirani, F. A. E. & Robinson, I. (1959), Proc. Roy. Soc. (Lond.), **A251**, 519.

Brinkman, H. W. (1923), Proc. Natl. Acad. Sci. (U.S.), **9**, 1.

Cartan, É. (1914), Ann. École Norm. Super., **31**, 263.

Cayley, A. (1860), Quart. J. of Pure and Appl. Math., **3**, 225 (see Collected Mathematical Papers, **4**, p. 447).

CAYLEY, A. (1869), Trans. Camb. Phil. Soc., **11** (**2**), 290 (see Collected Mathematical Papers, **7**, p. 66).
COXETER, H. S. M. (1936), Ann. of Math., **37**, 416.
CUNNINGHAM, E. (1910), Proc. Lond. Math. Soc., (**2**) **8**, 77.
DARBOUX, G. (1914), Leçons sur la Théorie Générale des Surfaces, Pt. 1, 2nd edn. (Gauthier-Villars, Paris).
DIRAC, P. A. M. (1936 a), Proc. Roy. Soc. (Lond.), **A155**, 447.
DIRAC, P. A. M. (1936 b), Ann. of Math., **37**, 429.
GEL'FAND, I. M., GRAEV, M. I. & VILENKIN, N. Ya. (1966), Generalized Functions, vol. 5: Integral Geometry and Representation Theory (Academic Press, New York).
GINDIKIN, S. G. (1983), Math. Intelligencer, **5**, no. 1, 27.
HEPNER, W. A. (1962), Nuovo Cim., **26**, 351.
HITCHIN, N. J. (1979), Math. Proc. Camb. Phil. Soc., **85**, 465.
HODGES, A. P. (1983), Proc. Roy. Soc. (Lond.) **A385**, 207; **A386**, 185.
HODGES, A. P. (1985), Proc. Roy. Soc. (Lond), **A 317**, 41, 375.
HODGES, A. P. & HUGGETT, S. A.; HUGHSTON, L. P.; PENROSE, R.; TOD, K. P.; WARD, R. S. (1980), Surv. High Eenergy Physics, **1**, 333, 313, 267, 299, 289.
HUGHSTON, L. P. (1979), Twistors and Particles; Lecture Notes in Physics 97 (Springer, Berlin).
KLEIN, F. (1870), Math. Annalen, **2**, 198.
KLEIN, F. (1926), Vorlesungen über höhere Geometrie (Springer, Berlin), pp. 80, 262.
KUIPER, N. H. (1949), Ann. of Math., **50**, 916.
LEE, T. D. & YANG, C. N. (1956), Phys. Rev., **104**, 254.
LIE, S. & SCHEFFERS (1896), Berührungstrsfn. 453.
MCLENNAN, Jr., J. A. (1956), Nuovo Cim., **10**, 1360.
MOUSSOURIS, J. (1983), Quantum Models of Space-Time Based on Recoupling Theory (D. Phil. thesis, Oxford).
MURAI, Y. (1953/4), Progr. Theoret. Phys., **9**, 147; **11**, 441.
MURAI, Y. (1958), Nucl. Phys., **6**, 489.
NEWMAN, E. T. (1976), Gen. Rel. Grav., **7**, 107.
NEWMAN, E. T. & PENROSE, R. (1968), Proc. Roy. Soc. (Lond.), **A305**, 175.
PENROSE, R. (1959), Proc. Camb. Phil. Soc., **55**, 137.
PENROSE, R. (1965 a), in Proceedings of the 1962 Conference on Relativistic Theories of Gravitation, Warsaw (Polish Acad. Sci. Warsaw).
PENROSE, R. (1965 b), Proc. Roy. Soc. (Lond.), **A284**, 159.
PENROSE, R. (1967), J. Math. Phys., **8**, 345.
PENROSE, R. (1968), Int. J. Theor. Phys., **1**, 61.
PENROSE, R. (1969 a), J. Math. Phys., **10**, 38.
PENROSE, R. (1969 b), in Quantum Theory and Beyond (ed. E. Bastin; Cambridge Univ. Press, Cambridge).
PENROSE, R. (1972 a), in General Relativity (ed. L. O'Raifeartaigh; Clarendon Press, Oxford).
PENROSE, R. (1972 b), in Magic without Magic (ed. J. R. Klauder; Freeman, San Francisco).
PENROSE, R. (1975 a), in Quantum Gravity, and Oxford Symposium (eds. C. J. Isham, R. Penrose & D. W. Sciama; Clarendon Press, Oxford).

PENROSE, R. (1975 b), in Quantum Theory and the Structures of Time and Space (eds. L. Castell, M. Dreischner & C. F. von Weizsacker; Carl Hanser Verlag, Munich).
PENROSE, R. (1976), Gen. Rel. Grav., **7**, 31, 171.
PENROSE, R. (1979), in Advances in Twistor Theory, Research Notes in Mathematics 37 (eds. L. P. Hughston & R. S. Ward; Pitman, San Francisco).
PENROSE, R. (1980), Gen. Rel. Grav., **12**, 225.
PENROSE, R. (1982 a), Bull. (New Ser.) Amer. Math. Soc., **8**, 427.
PENROSE, R. (1982 b), Proc. Roy. Soc. (Lond.), **A381**, 53.
PENROSE, R. & MACCALLUM, M. A. H. (1972), Phys. Repts., **6C**, 241.
PENROSE, R. & RINDLER, W. (1984), Spinors and Space-Time; vol. 1, Two-Spinor Calculus and Relativitic Fields (Cambridge Univ. Press, Cambridge).
PENROSE, R. & RINDLER, W. (1985), Spinors and Space-Time; vol. 2, Spinor and Twistor Methods in Space-Time Geometry (Cambridge Univ. Press, Cambridge).
PENROSE, R. & WARD, R. S. (1980), in General Relativity, One Hundred Years after the Birth of Albert Einstein (ed. A. Held; Plenum, New York).
PERJÉS, Z. (1975), Phys. Rev., **11D**, 2031.
PERJÉS, Z. (1979), Phys. Rev., **20D**, 1857.
PERJÉS, Z. & SPARLING, G. A. J. (1979), in Advances in Twistor Theory, Research Notes in Mathematics 37 (eds. L. P. Hughston & R. S. Ward; Pitman, San Francisco).
PLÜCKER, J. (1865), Phil. Trans. t., **155**, 725.
PLÜCKER, J. (1868/9), Neue Geometrie des Raumes gegründet auf die Betrachtung der geraden Linie als Raumdement (ed. F. Klein, Leipzig).
RADON, J. (1917), Ber. Verh. Sächs. Akad., **69**, 262.
REGGE, T. (1961), Nuovo Cim., **19**, 558.
ROBINSON, I. (1956), Report to the Eddington Group, Cambridge.
ROBINSON, I. (1961), J. Math. Phys., **2**, 290.
ROBINSON, I. & TRAUTMAN, A. (1962), Proc. Roy. Soc. (Lond.), **A265**, 463.
RUDBERG, H. (1958), dissertation, Univ. of Uppsala, Uppsala, Sweden.
RZEWUSKI, J. (1958), Nuovo Cim., **9**, 942.
SCHRÖDINGER, E. (1952), Science and Humanism (Cambridge Univ. Press, Cambridge).
SHAW, W. T. (1983), Proc. Roy. Soc. (Lond.), **A390**, 191.
SPARLING, G. A. J. (1975), in Quantum Gravity, an Oxford Symposium (eds. C. J. Isham, R. Penrose & D. W. Sciama; Clarendon Press, Oxford).
TOD, K. P. (1983), Proc. Roy. Soc. (Lond.), **A388**, 457.
WALKER, M. & PENROSE, R. (1970), Commun. Math. Phys., **18**, 265.
WARD, R. S. (1977), Phys. Lett., **61A**, 81.
WEIERSTRASS, K. (1866), Monatsberichte der K. P. Akademie, 612 (collected works, vol. 3).
WHITTAKER, E. T. (1903), Math. Annalen, **57**, 333.
WHEELER, J. A. (1962), Geometrodynamics (Acad. Press, New York).

177
NEW DIRECTIONS IN QUANTUM GRAVITY

Highlights in gravitation and cosmology. Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. Cambridge University Press, 1988

694 New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

NEW DIRECTIONS IN QUANTUM GRAVITY

Roger Penrose
Mathematical Institute, Oxford, U.K.

Quantum Gravity and its Infinities

There has been much work over the past thirty years or so, concerned with trying to discover how Nature is able to achieve unity and harmony in combining two seemingly incompatible collections of phenomena: those of the sub-microscopic world, described by quantum mechanics, and those of the large-scale world, described by general relativity. The essential need for such a quantum gravity theory had been emphasized particularly by Feynman (1958), at the 1957 Chapel Hill Conference. By means of a thought experiment involving a gravitational link in a quantum interference experiment, Feynman had argued that unless there were some new principle according to which quantum mechanics does not hold for masses approaching the Planck mass m_p ($\sim 10^{-5}$g) – the gravitational field, and hence space-time geometry itself must be subject to the rules of quantum theory: general relativity must indeed be quantized.

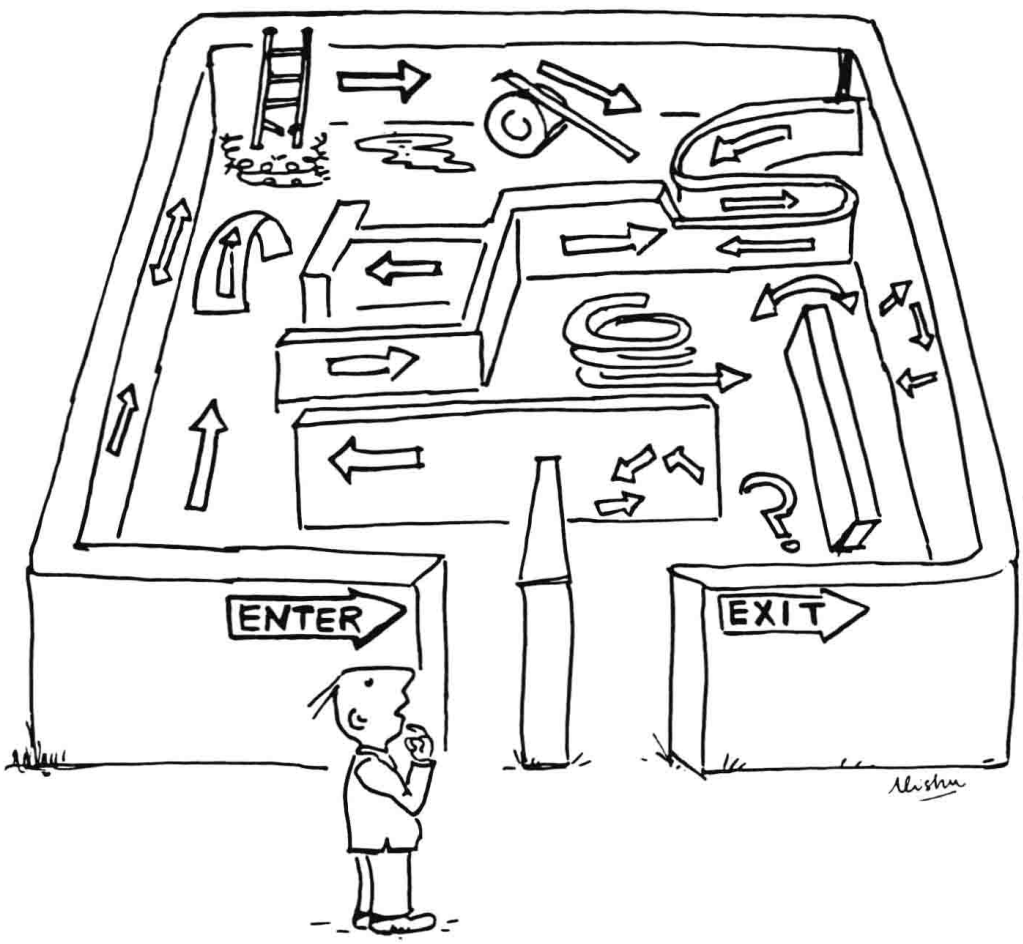
Numerous heroic attempts to quantize the Einstein theory followed (not least, those by Feynman himself); but these eventually foundered on the harsh rocks of non-renormalizability. Indeed, even mere renormalizability would not have been sufficient. One of the more "practical" original motivations for quantizing gravity had been to provide a cut-off at the order of the Planck length ℓ_p($\sim 10^{-33}$cm), so that the infinite renormalization parameters of standard quantum field theory could be replaced by finite quantities – quantities which might be, in principle calculable. This purpose would be self-defeating if the gravitational field had its own infinities.

Modifications of Einstein's Theory

This impasse led most workers in the field to explore possible modifications of Einstein's theory. One of the earliest of these was supergravity theory, according to which the graviton acquired a supersymmetric partner of spin 3/2. Other popular

New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

NEW DIRECTIONS



696 New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

modifications involved increasing the number of space-time dimensions, in the manner of Kaluza's and Klein's five-dimensional theory of unified gravitation and electromagnetism. Here one would contemplate a "space-time" of 4+N dimensions in which N of the dimensions were "compactified", i.e. curled up into some tiny ball - a space $\mathcal{C}$ of dimension N and linear size of the order of ℓ_p. The "Einstein vacuum equations" (i.e. Ricci-flatness) were normally assumed to hold for the (4+N)-dimensional space. Particularly popular had been an eleven-dimensional theory for which $\mathcal{C}$ was a certain "squashed 7-sphere". These higher-dimensional theories might also be combined with the ideas of supersymmetry. Moreover a different kind of modification of general relativity has sometimes been invoked, which involves replacing the standard (Hilbert) action of the theory (namely the scalar curvature) by something more complicated, so the field equations come out as something different from Ricci-flatness (for the vacuum). (This kind of modification also arises at the "effective" level, where one need not be seeking to modify general relativity as a classical theory, but merely to understand what modifications to the Einstein equations might arise as "quantum corrections" to the classical equations, at order $\hbar$.)

A good deal more radical are the now extremely fashionable superstring theories (see Green, Schwarz and Witten 1987). Here one attempts to replace the concept of a localized point particle by a loop (closed string) - or else an open string - whose history describes a two-dimensional surface. Motivational considerations require this surface to have a Lorentzian metric but, in accordance with the popular "Wick rotation" ideas in quantum field theory, the surface is taken, in mathematical discussions, to have a positive-definite (conformal) structure, i.e. to be a Riemann surface. This enables the theories to take advantage of the extremely elegant mathematics of such surfaces. Superstring theories developed from earlier appealing (due to Nambu) ideas which related to hadronic physics instead of to gravity (and which provided explanations for observational facts such as duality and the straightness of Regge trajectories). These earlier ideas had foundered on a seemingly fundamental difficulty: in order to have both gauge invariance and Lorentz invariance the space-time dimension has to be 26 rather than 4. It was later found (by Green and Schwarz) that if the theory is made supersymmetric, and so has a "fermionic" as well as "bosonic" aspect, then the "space-time dimension" could be 10 rather than 26. For the so-called "heterotic" string, moreover, this dimension seemingly contrives to be both 10 and 26 at the same time, according to whether, in a certain sense, it is left-travelling or right-travelling modes of oscillation of the string that are being considered. As one way of making sense of this kind of idea, the fundamental space of the theory is taken to be the string (-history)

New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

itself rather than the "space-time". Be that as it may, and whether or not one is concerned with this more exotic heterotic string, the extra space-time dimensions are taken to be "compactified" in accordance with the Kaluza-Klein type of idea. In addition to this is a gauge group structure, which may be either SO(32) or $E_8 \times E_8$. It is claimed that among the string oscillation modes are those corresponding a "graviton", and that a consistent quantum field theory for the string exists leading to, among other things, a consistent finite quantum gravity theory! Accordingly, the characteristic length scale of the theory is now taken ℓ_p rather than the strong interaction length ($\sim 10^{-13}$cm) of the original theory.

I may say that I do not, myself, regard these ideas as being at all satisfying as they stand. The geometry of the situation, particularly for the heterotic string, is rather obscure. Even where the "space-time" that arises makes clear and consistent sense, I find it hard to believe that the "compactification" mechanism of the Kaluza-Klein type of idea will actually work. There are reasons to believe that, in the classical theory, such compactified dimensions will be generically unstable (as some recent work in collaboration with Arvind Borde, in which the techniques of classical singularity theorems are invoked, seems to indicate). One would need to depend heavily on the quantum theory to have any hope that the classical singularities resulting from this instability might be avoided. Moreover, there is no significant evidence that quantum theory actually does lead to the avoidance of space-time singularities. My own guess is that it does not. In any case it is pure speculation that such a Kaluza-Klein type of mechanism would actually be operative and that a seeming four-dimensional space-time would actually arise out of all this, with a metric that at the classical level, satisfies the four-dimensional Einstein equations (rather than that for 10 or 26 dimensions). In any case, it is quite obscure to me how a consistent-looking curved space-time actually arises, starting from the string theory picture. In my opinion, far from providing a consistent quantum gravity theory, the present superstring ideas have left us a very great deal farther from such a theory than we were before!

The Problem of Time-Asymmetry

There is another important issue that I wish to address, and which has been left totally untouched not only by superstring theory, but by all other serious approaches to quantum gravity that I have seen (including my own attempts via twistor theory). This is the issue of time-asymmetry in space-time singularity structure (cf. Penrose 1979, 1981). One of the original motivations lying behind the desire for a quantum gravity theory was that the existence of singularities is a necessary feature of the classical theory. One

 New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

would seem to need a quantized space-time theory in order that the physics "at" a space-time singularity can be understood, and so that the presence of an actual singularity in the mathematical description might be avoided. Our understandings of cosmogeny, whether or not they invoke the fashionable ideas of inflation, have reached the impasse of the initial big-bang singularity. We need to know why the detailed structure of this initial singularity is so close to having a Friedman-Robertson-Walker form. There must have been an enormous constraint on the singularity, since it is very far from being "generic". For example, for a universe that develops to one containing 10^{80} baryons, the constraint on the initial state restricts us to a volume of phase-space that is no greater than one part in $10^{10^{123}}$ of the total. The volume could be even smaller if inflation is present!

One can begin to see how great this constraint must be when we consider a collapsing universe containing irregularities. Black holes will form and their singularities are expected to unite to form the final big-crunch singularity. The entropy can be estimated by use of the Bekenstein-Hawking formula for black-hole entropy:

$$S_{BH} = \frac{A}{4} \cdot \frac{kc}{G\hbar}^3$$

(where k is Boltzmann's constant, etc., and A is the surface area of the horizon). The constraint that evidently applied to the big bang does not (in time-reversed form) apply to the big crunch, nor to the singularities in black holes. it is this time-asymmetry in singularity structure that gives us, among other things, the second law of thermodynamics. For the second law to hold, the initial singularity must have been highly constrained and and have very low entropy, whereas this cannot be true of final singularities, which must have very high entropy. The nature of this constraint on initial singularities appears to be that the Weyl curvature goes to zero as the singularity is approached from the future direction. I refer to this constraint as the Weyl curvature hypothesis (WCH).

We note that this hypothesis is grossly time-asymmetric, since, in accordance with the second law, the constraint cannot apply to final singularities. Since one of the tasks of quantum gravity theory is to explain the structure of space-time singularities, it would seem that WCH, or something very like it, must be one of its implications. But WCH is time-asymmetric and this seems to imply that quantum gravity theory must therefore be time-asymmetric also. (The alternative would seem to be some sort of symmetry-breaking mechanism, but it is very hard to see how such a mechanism could be used to derive the second law, when in the usual way that it is applied, it

New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

actually depends upon that law! For an alternative point of view, see Hawking 1987.) However, general relativity is a time-symmetric theory. So is superstring theory, twistor theory, and also all the gauge theories that go into the physics of particle theory and inflationary universe models.

Time Asymmetry in Quantum Probabilities

Moreover, quantum theory itself is also time-symmetric in the form that it is employed. The part of quantum theory that is always used is the unitary-evolution part of quantum theory. There is also another part, namely that which is employed when probability amplitudes are converted to probability and we run into the issue of quantum measurements and the "collapse of the wave function". I believe that a clear case can be made that this second part is indeed independent of the first. Moreover, it is time-asymmetric (cf. Penrose 1987). We cannot use the procedure in the reverse direction in time to obtain correct probabilities.

For example, consider a photon source aimed at a photo-cell between which is placed a half-silvered mirror tilted at 45° to the line from source to cell. Let the system be surrounded by walls. Now, if we are given that the source has emitted a photon, the quantum-mechanical calculation correctly gives 50% probability that the photon reaches the photo-cell and 50% probability that it strikes the wall. However, given that the photo-cell registers receipt of the photon, the reversed-time quantum-mechanical calculation incorrectly gives 50% probability that the photon left the source and 50% probability that it came from the wall, the correct probabilities being, in fact, close to 100% and 0%, respectively. (See Penrose 1987 for further discussion.)

Towards a Physical Theory of Wave Function Collapse

Thus, the "collapse of the wave function" part of quantum mechanics is indeed a time-asymmetric process. My own point of view (Penrose 1981, 1987) is that this process is something physically real, independent of our subjective experiences, for which a new theory will be needed. In my opinion, it is our task to find a time-asymmetric quantum gravity theory which will have, among its consequences, both a resolution of the quantum measurement problem and an explanation of WCH. Thus, it is my view it is not general relativity that, as a classical theory, must be changed, but it is quantum theory that needs overhaul! More correctly, it is both theories that must be prepared for modification. Perhaps this is somewhat in line with Feynman's "new principle" mentioned earlier.

 New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

Preliminary calculations indicate that it is not unreasonable to expect that an objective wave-function collapse can occur at a level (the "one-longitudinal-graviton level") at which gravitation begins to be involved in a quantum process. More work needs to be done on this.

As regards where to look in order to find such a theory, my own predelictions are to examine some form of twistor theory (cf. Huggett and Tod 1985, Penrose and Rindler 1986) into which a time-asymmetry has been injected. Because of the existence of EPR-type effects which violate Bell's inequalities, one will need some description which is essentially non-local. Twistor theory treats space-time in an essentially non-local way. It also fits in well with some of the more attractive aspects of string theory (Shaw 1985, Hughston and Shaw 1987), but in my view it will be necessary to get a string theory down into four space-time dimensions. This has also been a recent trend among some string theorists. I believe that there is some reasonable hope that all the appropriate ideas can eventually be brought satisfactorily together in a comprehensive theory. Indeed, a proposal due to M.A. Singer and A.P. Hodges now unifies string-theoretic and twistor-theoretic ideas in an exciting new way!

References

Green, M.B., Schwarz, J.H. and Witten, E. (1987) Superstring Theory (Cambridge Univ. Press: Cambridge)

Feynman, R.P. (1958) in Proceedings of the Chapel Hill Conference, 1957 Revs. Mod. Phys.

Hawking, S.W. (1987) Quantum Cosmology, in 300 Years of Gravity eds. S.W. Hawking and W. Israel (Cambridge Univ. Press: Cambridge)

Huggett, S.A. and Tod K.P. (1985) An Introduction to Twistor Theory (Cambridge Univ. Press: Cambridge)

Hughston, L.P. and Shaw, W.T. (1987) Classical strings in ten dimensions, Proc. Roy. Soc. (Lond.) to appear.

Penrose, R. (1979) Singularities and Time Asymmetry, in General Relativity: an Einstein Centenary eds. S.W. Hawking and W. Israel (Cambridge Univ. Press: Cambridge)

Penrose, R. 81) Time-Asymmetry and Quantum Gravity, in Quantum Gravity 2, eds. C.J. Isham, R. Penrose and D.W. Sciama (Oxford Univ. Press, Oxford)

New directions in quantum gravity. *Highlights in gravitation and cosmology.* Eds B. R. Iyer, A. Kembhavi, J. V. Narklikar, and C. V. Vishveshwara (with illustrations by C. V. Vishveshwara), 175–82. 1988

Penrose, R. (1987) Newton, Quantum Theory and Reality, in 300 Years of Gravitation eds. S.W. Hawking and W. Israel (Cambridge Univ. Press: Cambridge)

Penrose, R. and Rindler, W. (1968) Spinors and Space-Time, Vol. 2 Cambridge Univ. Press: Cambridge)

Shaw, W.T. (1985) Twistors, minimal surfaces and strings, Class, Quantum Grav. 2, L113-119.

178
PRETZEL TWISTOR SPACES

Twistor Newsletter **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, Longman, 1990

This interesting idea, due to Michael Singer and Andrew Hodges, is developed a little further here; see also Chapter 186.

 Pretzel twistor spaces. *Twistor Newsletter* **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

§I.4.22 Pretzel twistor spaces *by R.Penrose* (TN 26, March 1988)

In the previous article §I.4.21, Singer has provided a very appealing suggestion for combining some of the attractive features of string theory/conformal field theory with twistor theory. As an analogue of the (bounded) Riemann surfaces of string theory/conformal field theory (pretzels), one imagines a complex manifold—say of 3 dimensions or 4 dimensions, according to whether we are considering projective or non-projective twistors—with a boundary consisting of a number of copies of $\mathbb{PN}$ or $\mathbb{N}$, as the case may be, see figure 1. (See also Tsou S.T.'s following article §I.4.23 for a suggestion whereby these $\mathbb{PN}$'s shrink down to lines.) The most appropriate formulation of this idea remains somewhat unclear as of now, but we might think in terms of some sort of analogue of the Fock spaces that in string theory/conformal field theory are to be specified at each connected portion of the boundary—say, at each '*end*'. Then amplitudes could be obtained when elements of the appropriate Fock spaces are specified at each end.

From the point of view of twistor theory, a Fock space (in its normal physical interpretation) might seem inappropriate. For reasons of wishing to tie up this idea with twistor diagram theory—ideas due to Hodges (see TN 26 and next chapter)—and of exploiting the relevant crossing symmetry, spin-statistics, duality, etc., we might prefer to think of the 'ends' as corresponding to *individual* particles, rather than states involving an

indefinite number of particles as described by Fock space elements.

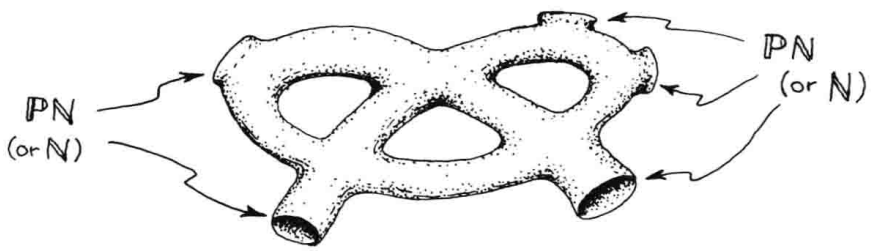


Figure 1

The twistorial description of Fock space elements would be by certain elements of

$$\mathbb{C} \oplus H^1(\mathbb{T}^+,\mathcal{O}) \oplus H^2(\mathbb{T}^+\times\mathbb{T}^+,\mathcal{O}) \oplus H^3(\mathbb{T}^+\times\mathbb{T}^+\times\mathbb{T}^+,\mathcal{O}) \oplus \cdots$$

i.e. by collections of 'twistor functions'

$$f_0,\ f_1(Z^\alpha),\ f_2(Z^\alpha,Y^\alpha),\ f_3(Z^\alpha,Y^\alpha,X^\alpha),\cdots$$

which are, respectively '0-functions', 1-functions, 2-functions, 3-functions, etc., where the n-function f_n is the twistor wave function for a state consisting of n massless particles. (The '0-function' f_0 would simply be an element of $\mathbb{C}$, defining the amplitude for zero particles.) Instead of this, we might well prefer to think of f_0, f_1, f_2, f_3, $\cdots$ as referring to the functions of several twistors which come up in the *twistor particle programme*. Thus, to a 'first approximation', f_1 might describe the amplitude for the state of a massless particle, f_2 the state of a lepton (?), f_3 the state of an 'ordinary' hadron (??), f_4, etc., for more 'exotic' hadrons (???) (and f_0 for a 'nothing'), but an actual massive particle might be imagined to have small contributions all indefinitely far along the sequence. We expect f_1 to be a 1-function[1] and suspect that f_2 might perhaps be a *relative* 1-function (or 2-function?) relative to some sort of diagonal locus in $\mathbb{T}\otimes\mathbb{T}$, conformal invariance being broken at this stage (so mass and $I_{\alpha\beta}$ come in). Perhaps *all* the f_i are relative 1-functions (or 2-functions) or something (whence $f_0\equiv 0$). Perhaps, as with strings, the higher fs describe higher 'modes', where oscillation of strings are replaced by deformations of **PN**. Perhaps, in accordance with an idea floated by G.B.Segal, they refer to different jet bundles over (ℙ)ℕ.

[1] I use the term 'n-function' to denote an element of n^{th} (holomorphic) sheaf cohomology.

 Pretzel twistor spaces. *Twistor Newsletter* **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

All this is extremely vague and speculative, as yet. Let us backtrack a little and try to see whether there are, indeed, any interesting 'pretzel' twistor spaces. As with string/conformal field theories, *closed* pretzel spaces are of particular interest. The simplest way to construct such a twistor space (considering the case of pretzelized $\underline{\mathbb{P}\mathbb{T}}$s only) would seem to be the following. Consider first the two limiting projective motions of $\mathbb{P}\mathbb{T}$:

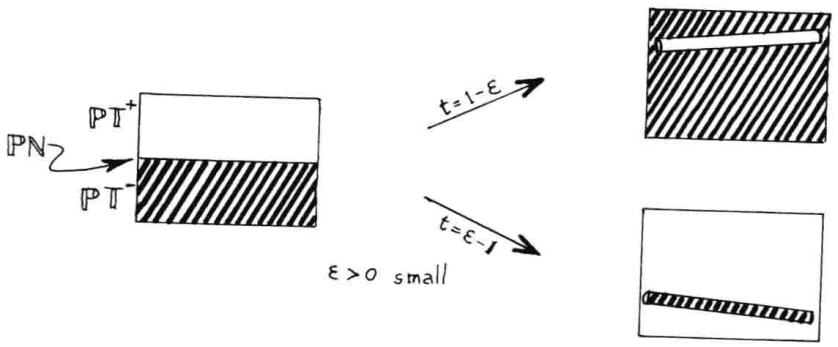


Figure 2

These may be achieved by

$$Z^{\alpha} = \begin{bmatrix} \omega^0 \\ \omega^1 \\ \pi_{0'} \\ \pi_{1'} \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ t & 0 & 1 & 0 \\ 0 & t & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega^0 \\ \omega^1 \\ \pi_{0'} \\ \pi_{1'} \end{bmatrix} = Z_t^{\alpha}$$

$$Z_t^{\alpha} = \begin{bmatrix} \omega_t^1 \\ \omega_t^0 \\ \pi_{t\,0'} \\ \pi_{t\,1'} \end{bmatrix}$$

with $t \in (-1,1)$, the two limiting cases being $t \to 1$, $t \to -1$, respectively.

The 5-surface into which $\mathbb{P}\mathcal{N}$ is carried (i.e. given by $Z_{-t}^{\alpha} \bar{Z}_{-t\alpha} = 0$) is another copy of $\mathbb{P}\mathcal{N}$ (identical with it as a CR-manifold)—call it $\mathbb{P}\mathcal{N}_t$. As $t \to 1$, this 5-surface closes in on the line $L^+ \subset \mathbb{P}\mathbb{T}^+$ given by

$$Z^{\alpha} = \begin{bmatrix} \lambda \\ \mu \\ \lambda \\ \mu \end{bmatrix},$$

and as $t \to -1$ it closes in on $L^- \subset \mathbb{P}\mathbb{T}^-$, given by $Z^{\alpha} = \begin{bmatrix} \lambda \\ \mu \\ -\lambda \\ -\mu \end{bmatrix}$.

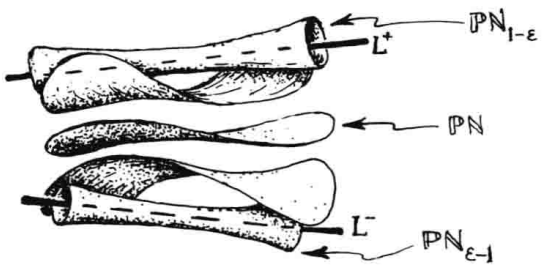


Figure 3

Note that if we define $|Z|^2 = |\omega^0|^2 + |\omega^1|^2 + |\pi_{0'}|^2 + |\pi_{1'}|^2$, then $|Z_t|^2 = (1+t^2)|Z|^2 + 2tZ^\alpha \bar{Z}_\alpha$ and $Z_t^\alpha \bar{Z}_{t\alpha} = (1+t^2)Z^\alpha \bar{Z}_\alpha + 2t|Z|^2$. Hence $\mathbb{PN}_t$ is defined by

$$Z^\alpha \bar{Z}_\alpha : |Z|^2 = 2t{:}1+t^2.$$

$\mathbb{PN}_{1-\epsilon}$ (resp. $\mathbb{PN}_{\epsilon-1}$) is the boundary of a tubular neighbourhood of L^+ (resp. L^-).

The simplest pretzel space $\mathcal{P}_1$ is given by identifying Z^α with Z_t^α for some fixed T in (0,1) (and hence, up to proportionality, Z^α is also identified with Z_t^α where $t = \tanh n\tau$, for all $n \in \mathbb{Z}$). This can also be achieved by drilling out the regions (inside) above $\mathbb{PN}_{t'}$ and below (inside) $\mathbb{PN}_{-t'}$ (where $t' = \tanh \frac{1}{2}\tau$) and identifying the boundaries. There is actually some freedom in how such an identification is made, and this allows *different* pretzel twistor spaces to be built in this way.

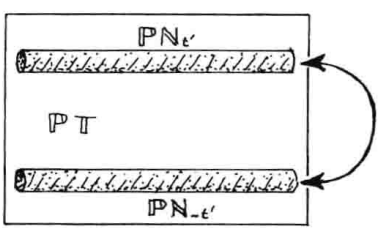


Figure 4

The spaces $\mathcal{P}_1$ are the analogues of *toruses*, i.e. of Riemann surfaces of genus 1. For higher genus (genus g), we can carry out several such identifications simultaneously to give a '$\mathbb{CP}^3$ with g handles'. Perhaps easier to visualize, but equivalent, is to take *two* copies of $\mathbb{CP}^3$ and to identify across from one to the other.

 Pretzel twistor spaces. *Twistor Newsletter* **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

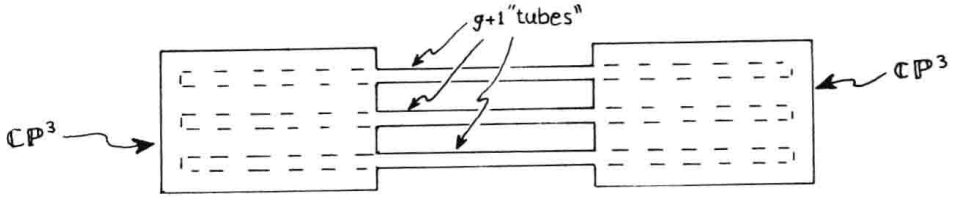


Figure 5

We could also do

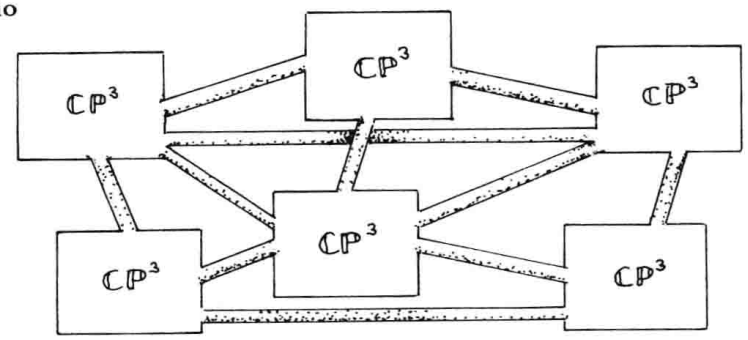


Figure 6

etc., but this gives us no more generality. (To see this, clip all the tubes necessary to make a tree; then fit all the $\mathbf{CP}^3$s inside *one* of them and we are back with the case of 'handles' as before, when we re-glue where we had clipped.)

The analogies between these spaces and *Riemann surfaces* are quite striking. In particular, we can ask for the *dimension* m of *moduli space* for pretzel twistor spaces of genus p (i.e., how many complex parameters are needed to characterize a $\mathcal{P}_g$, where a $\mathcal{P}_g$ is a $\mathbf{CP}^3$ with g 'handles' of this type). (I do not know whether this 'genus' corresponds to something standard in algebraic geometry.) Recall that for Riemann surfaces the answer (due to Riemann) is $3g-3$, except when $g=0$ or 1—where for $g=0$ the answer is 0 and for $g=1$ it is 1. For $\mathcal{P}_g$ the answer turns out to be precisely 5 *times* as large—with the *single exception* that for $\mathcal{P}_1$, $m=3$ (instead of 5). The proof is similar to that for Riemann surfaces. Think first of a *labelled* $\mathbf{CP}^3$. The labelling can be achieved by specifying 5 points in general position on $\mathbf{CP}^3$.

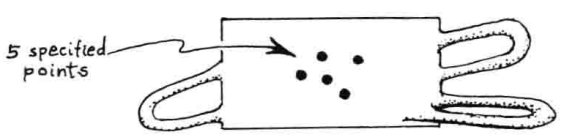


Figure 7

There are 15 complex degrees of freedom in the specification of each pipe. (Or do it with with 2 copies of $\mathbf{CP}^3$ as in Figure 5 which may be a little easier to visualize.) The 15 comes from the size of the projective group on $\mathbf{CP}^3$ ($15=4^2-1$). We have $15g$ for the number of parameters needed to define a *labelled* $\mathcal{P}_g$. We have to factor out by the freedom in doing the labelling—which is 15 parameter's worth *unless*, in the *generic* case, there are continuous (holomorphic) symmetries of $\mathcal{P}_g$. We do *not* factor out by the motions of the labelling corresponding to a symmetry. Suppose that there are d dimensions of symmetries. Then we get

$$m=15g-(15-d).$$

When $g=0$ then clearly $d=15$, so $m=0$. When $g=1$ then a *direct* argument shows that $m=3$ (whence $d=3$ in the generic case). (This argument is: $\mathcal{P}_1$ is defined by $Z^\alpha \equiv T^\alpha{}_\beta Z^\beta$, up to proportionality, for some fixed $T^\alpha{}_\beta$. The ratios of the eigenvalues of $T^\alpha{}_\beta$ give the moduli of $\mathcal{P}_1$.) It is not hard to see that $d=0$ whenever $g\geq 1$ (as with Riemann surfaces), so $m=15g-15$ in these cases.

I should remark that there is a subtlety involved in the glueing of the pipes together ('handles'). When we think of $\mathbb{PN}_{1-\epsilon}$ surrounding L^+, we think of it as an $S^2\times S^3$ for which the S^3's surround the various points of L^+) and where the different points of L^+ give the S^2:

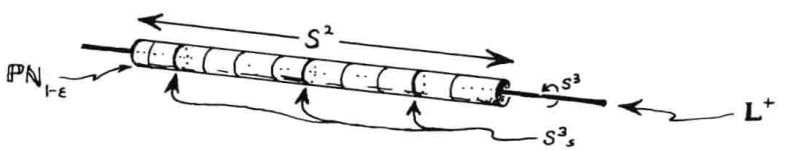
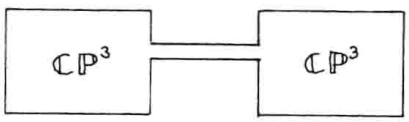


Figure 8

If we glue this to $\mathbb{PN}_{\epsilon-1}$ we must do so in such a way that the small S^3s of $\mathbb{PN}_{1-\epsilon}$ are stretched the length of $\mathbb{PN}_{1-\epsilon}$,

Figure 9

and vice versa. This is possible because of a topological relation which I write symbolically as $S^2\times S^3\simeq S^3\times S^2$, each side being a circle bundle over $S^2\times S^2$, but where the twist can be transferred from being over the second S^2 factor to over the first. (This is the Clifford-Hopf bundle of S^1 over S^2 to give S^3.) If we take planes through L^-, they sweep out one family of S^3s on each $\mathbb{PN}_t$ ('small S^3s' when $t=1-\epsilon$, and running the 'length' of $\mathbb{PN}_t$ when $t=\epsilon-1$); planes through L^+ sweep out the other family of S^3s ('long' when $t=1-\epsilon$ and 'small' when $t=\epsilon-1$). The relation '$S^2\times S^3\simeq S^3\times S^2$' can also be seen by examining pairs of orthogonal unit vectors at the origin of $\mathbb{R}^4$. Fix attention on one vector: it sweeps out an S^3 while the other gives a (trivial) bundle over it (trivial since S^3 is parallelizable). Then think of the vectors in the other order.



$\mathbb{CP}^3$, *so long as tube is not pinched*

Figure 10

If we consider a *thin* tube, where the tube narrows down to nothing, we get in the limit the space (considered by twistorians in connection with null lines on $\mathbb{CI}$, and by Donaldson in a context similar to the present one) which is two $\mathbb{CP}^3$s joined along a quadric, which is a blown-up line in each, and identified with generator systems reversed:

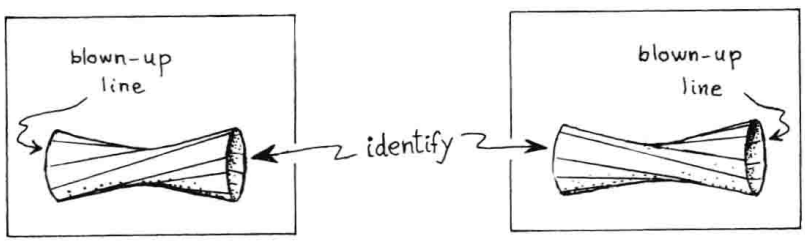


Figure 11

The topology of $\mathcal{P}_g$ is an S^2 bundle over S^4 with g handles ($S^3\times\mathbb{R}$ handles).

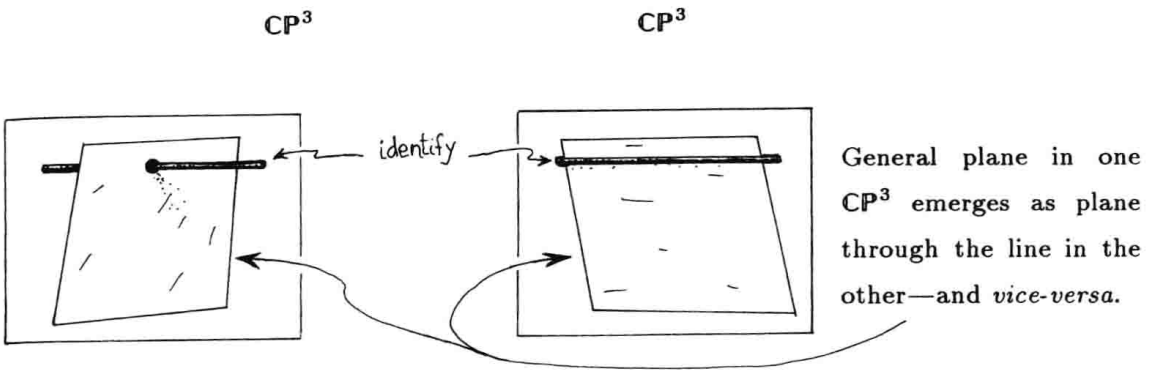
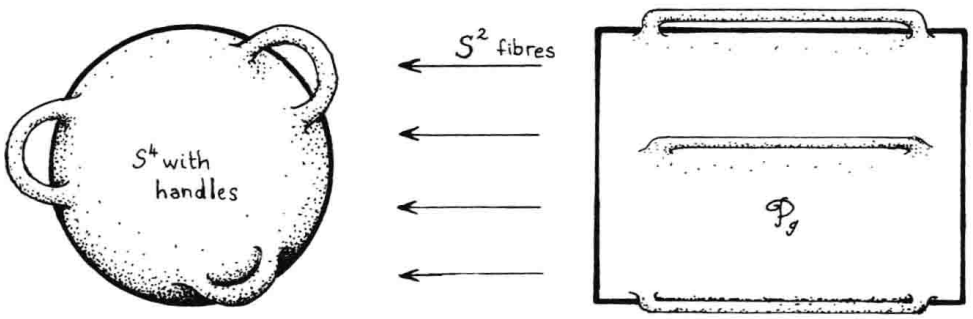

$\mathbb{CP}^3$
$\mathbb{CP}^3$
identify
General plane in one $\mathbb{CP}^3$ emerges as plane through the line in the other—and *vice-versa.*

Figure 12



S^4 with handles $\mathbb{CP}^3$ with tubes

Figure 13

These spaces $\mathcal{P}_g$ are ones where there are $\mathbb{CP}^1$s with neighbourhoods which are identical with portions of $\mathbb{CP}^3$—and are twistor spaces of *conformally flat* 4-spaces.

Much more can be said. Of course, in line with non-linear graviton constructions we shall want to *deform* these spaces; also to deform the CR-structure of $\mathbb{PN}$, etc. Work in progress. Thanks to M.A.Singer, A.P.Hodges, G.B.Segal, T.S.Tsou, E.Dunne, R.Baston, and S.K.Donaldson.

179
TWISTORS AND STATE-VECTOR REDUCTION

Twistor Newsletter **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

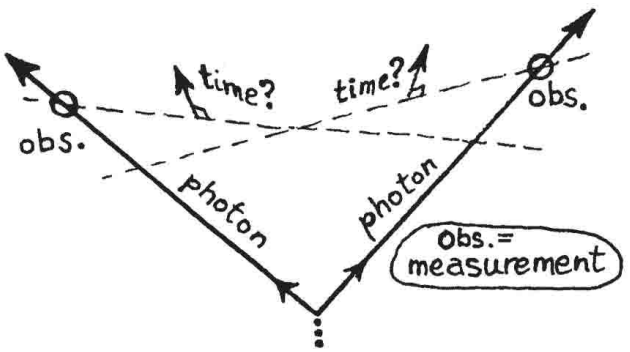
Might twistor theory conceivably ever shed any light on quantum state reduction?

 Twistors and state-vector reduction. *Twistor Newsletter* **26** (1988). Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.4.30 **Twistors and State-Vector Reduction** *by R. Penrose* (TN 26, March 1988)

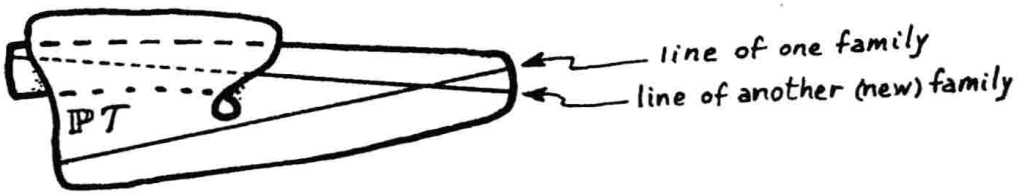
One of the most puzzling features of the procedure of the "reduction of the state vector" in quantum theory is that there seems to be no particular "moment" at which

it happens. Yet, in the real world of our experience, such discontinuous changes in the quantum-mechanical description of the world have to be considered to take place, generally, between two measurements. One might have thought that any such violently discontinuous change in the state of the world—if it were a *real* effect—ought to be more noticeable, as to *when* it actually takes place! Also, since the Schrödinger equation is such a nice smooth analytic thing, it seems odd that Nature should choose to execute such violent discontinuous jumps from time to time. Perhaps most baffling is the non-local and seeming relativity-conflicting behaviour in EPR-type (Clauser-Aspect) experiments. When space-like-separated measurements take place, there is a conflict between the apparent time-ordering of the "reductions" due to these two measurements. "*When*" do these reductions "*actually*" take place?



It is conceivable that a twistor-type viewpoint could provide some sort of resolution of this puzzle. Suppose that reduction is a gravitational effect (cf. Penrose 1989, for example, and the references contained therein) and that the space-time is described twistorially. Of course we need to have solved the problem of representing *general* space-times, not merely anti-self-dual ones, for this to work (beyond the googly!) and it may well be that space-time points are not simply "lines" in twistor space. But for the sake of descriptiveness (only), let's take lines. Now suppose that, with some measurement, the twistor space $\mathbb{P}\mathcal{T}$ becomes sufficiently "curved" that the original family of holomorphic lines, representing space-time points, peters out—and we must switch to a *new* family of such lines in order to keep going. Somehow the geometry of space-time seems to jump—yet in the "actuality" of the twistor space there is *no* jump—just a (necessary) shift in viewpoint. Anyway, it's a thought!

716 Twistors and state-vector reduction. *Twistor Newsletter* **26** (1988). Also published in *Further Advances in Twistor Theory,* Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001



References

Penrose, R. (1989) *The Emperor's New Mind,* Oxford University Press, Oxford.

180
ASPECTS OF QUASI-LOCAL ANGULAR MOMENTUM

Mathematics and general relativity (1986), *Contemporary Mathematics* **71**. Ed. J. Isenberg, 1–8. American Mathematical Society, Providence, 1988

This was over 10 years before convincing observational evidence of the presence of a positive cosmological constant was at hand. Could the comments in the final paragraph (on p. [7]) have any relevance to these new findings?

 Aspects of quasi-local angular momentum. *Mathematics and general relativity* (1986), *Contemporary Mathematics* **71**. Ed. J. Isenberg, 1–8. 1988

Aspects of Quasi-Local Angular Momentum

by

Roger Penrose
Mathematical Institute
Oxford, U.K.

and

Institute for Theoretical Physics
Santa Barbara, California, U.S.A.

I wish to make some remarks concerning the expression that I suggested a few years ago (Penrose 1982; and as subsequently slightly modified, c.f. Penrose and Rindler 1986), which provides a possible definition for the total energy-momentum and angular momentum surrounded by a spacelike 2-surface $\mathcal{S}$, with S^2 topology, in a general-relativistic (four-dimensional) space-time $\mathcal{M}$. The definition is quasi-local in the sense that it refers only to the geometry of $\mathcal{S}$ and the extrinsic curvature quantities for its embedding in $\mathcal{M}$. The definition does not, as things stand, enable one to pick out the family of 4-momentum components from the complete complex of quantities that are defined. Indeed, in the general case, there is, as yet, no unambiguous choice for the definition of even a scalar rest-mass in terms of this complex of quantities. The difficulty here is that there are sometimes different invariant definitions for this rest-mass, and it could be simply a matter of making a "correct" choice from amonst the various possibilities. The prospect of separating out a 4-momentum from the angular momentum in an invariant way is obstructed in a more fundamental way, however. It is the main purpose of this note is to explain why this must be so.

I should first recall the definition of the complex $A_{\alpha\beta}(=A_{\beta\alpha})$ of these quantities:

$$A_{\alpha\beta}Z^{\alpha}Z^{\beta} = \frac{-1}{16\pi}\oint \eta\, R_{abCC'DD'}\omega^{C}\omega^{D}\epsilon^{C'D'}\, dx^{a} \wedge dx^{b}$$

integrated over a spacelike 2-surface $\mathcal{S}$ ($\approx S^2$), where Z^α stands for the 2-surface twistor defined by the field ω^A on $\mathcal{S}$, subject to the <u>2-surface twistor equation</u>

$$\eth' \omega^0 = \sigma' \omega^1, \qquad \eth \omega^1 = \sigma \omega^0$$

(ω^0 and ω^1 being the components of ω^A referred to a spin-frame o^A, ι^A whose flagpoles are orthogonal to $\mathcal{S}$), and where η is proportional, over the surface $\mathcal{S}$ to the determinant

$$\begin{vmatrix} \omega^0_1 & \omega^0_2 & \omega^0_3 & \omega^0_4 \\ \omega^1_1 & \omega^1_2 & \omega^1_3 & \omega^1_4 \\ \eth\omega^0_1 & \eth\omega^0_2 & \eth\omega^0_3 & \eth\omega^0_4 \\ \eth'\omega^1_1 & \eth'\omega^1_2 & \eth'\omega^1_3 & \eth'\omega^1_4 \end{vmatrix}$$

for independent solutions $\overset{A}{\omega}_1, \ldots, \overset{A}{\omega}_4$ of the 2-surface twistor equation. (See Penrose and Rindler 1984 for the notation used here.) In the weak-field limit of general relativity, or at null or spacelike infinity (when $\mathcal{M}$ is asymptotically flat), there is a Hermiticity condition reducing the ten complex components of $A_{\alpha\beta}$ to ten real ones, namely:

$$A_{\alpha\gamma} \bar{I}^{\beta\gamma} = \bar{A}^{\beta\gamma} I_{\alpha\gamma}$$

(Penrose 1982, Penrose and Rindler 1986, Shaw 1983), where the <u>infinity twistor</u> $I_{\alpha\beta}$, subject to

 Aspects of quasi-local angular momentum. *Mathematics and general relativity* (1986), *Contemporary Mathematics* **71**. Ed. J. Isenberg, 1–8. 1988

$$\tilde{I}^{\alpha\beta} = \frac{1}{2}\,\epsilon^{\alpha\beta\gamma\delta} I_{\gamma\delta}, \qquad I_{\alpha\beta} = -I_{\beta\alpha}, \qquad \det(I_{\alpha\beta}) = 0\,,$$

has a (flat-space) definition

$$I_{\alpha\beta} = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon^{A'B'} \end{pmatrix}$$

in the description whereby any k-valent twistor has 2^k spinor parts, each of which is a spinor field. These spinor fields are related to one another via certain differential equations. In the case of Z^α, the spinor parts are

$$Z^\alpha = (\omega^A, \pi_{A'})$$

where, in Minkowski space $\mathbb{M}$, we have

$$\nabla_{AA'}\omega^B = -i\epsilon_A{}^B \pi_{A'}, \qquad \nabla_{AA'}\pi_{B'} = 0,$$

and the relations between the spinor parts of a polyvalent twistor are generated from these (Penrose and Rindler 1986). For the Hermiticity condition on $A_{\alpha\beta}$ we require not only an infinity twistor but, in addition, a <u>complex-conjugacy</u> operator

$$Z^\alpha \to \bar{Z}_\alpha$$

which is also a duality operation. This is equivalent to providing each twistor Z^α with its "twistor norm"

$$Z^\alpha \bar{Z}_\alpha = \omega^A \bar{\pi}_A + \pi_{A'} \bar{\omega}^{A'}$$

which is Hermitian, with signature (++--).

In conformally flat space-time we can still use a description of twistors in terms of spinor parts, but where we now have

$$\nabla_{AA'}\omega^B = -i\epsilon_A{}^B \pi_{A'}, \qquad \nabla_{AA'}\pi_{B'} = -iP_{ABA'B'}\omega^B$$

where

$$P_{ab} = \frac{1}{12} R\, g_{ab} - \frac{1}{2} R_{ab}.$$

The expression for the twistor norm is the same as for $\mathcal{M}$, as given above.

In deSitter space-time, with cosmological constant $\lambda = 6\Lambda$, we have

$$\nabla_{AA'}\omega^{B} = -i\varepsilon_{A}{}^{B}\,\pi_{A'}, \qquad \nabla_{AA'}\pi_{B'} = -i\Lambda\;\varepsilon_{A'B'}\,\omega_{A}$$

so ω^{A} and $\pi_{A'}$ are essentially on an equal footing. The spinor-part representation of the deSitter-space infinity twistor is

$$I_{\alpha\beta} = \begin{pmatrix} \Lambda\varepsilon_{AB} & 0 \\ 0 & \varepsilon^{A'B'} \end{pmatrix} .$$

As in Minkowski space, its complex conjugate is its dual (i.e. "twistor-real"), but now

$$\det(I_{\alpha\beta}) = \Lambda$$

and we have

$$I_{\alpha\gamma}I^{\beta\gamma} = \Lambda\,\delta_{\alpha}^{\beta} \quad .$$

In a general curved space-time $\mathcal{M}$, one cannot provide an integrable concept of twistor in terms of such spinor fields, but restricted to $\mathcal{S}$ we have

$$\eth\omega^{0} = -i\rho'\pi_{1'}, \qquad \eth'\omega^{1} = -i\rho\pi_{0'} \quad .$$

At any one point of $\mathcal{S}$, one can write down the previous expression for the "twistor norm", namely

$$\omega^{0}\bar{\pi}_{0} + \omega^{1}\bar{\pi}_{1} + \pi_{0'}\bar{\omega}^{0'} + \pi_{1'}\bar{\omega}^{1'}$$

but this is generally not constant over $\mathcal{S}$. The condition for it to be constant for all 2-surface twistors Z^{α} is that $\mathcal{S}$ be uncontorted (e.g. cf. Jeffryes 1984), i.e. that $\mathcal{S}$ can be embedded in some conformally flat real space-time in such a way that both its intrinsic geometry (induced

Aspects of quasi-local angular momentum. *Mathematics and general relativity* (1986), *Contemporary Mathematics* **71**. Ed. J. Isenberg, 1–8. 1988

metric) and its extrinsic curvature quantities $(\rho,\rho',\sigma,\sigma')$ are unaltered from their values in $\mathcal{M}$. In these uncontorted cases, this provides an unambiguous definition for $Z^\alpha\bar{Z}_\alpha$ and, therefore, for the complex-conjugation map $Z^\alpha \to \bar{Z}_\alpha$. When $\mathcal{S}$ is contorted, on the other hand, we seem to need some means of averaging this local twsitor norm expression over $\mathcal{S}$. There appear to be various inequivalent ways of achieving this averaging (cf. Penrose 1982), but none has been suggested which is conformally invariant. It is not altogether clear whether such conformal invariance is desirable since general relativity is not, after all, a conformally invariant theory, but the definition of a 2-surface twistor is a conformally invariant one.

With regard to $I_{\alpha\beta}$, we certainly would not expect conformal invariance. In $\mathcal{M}$, we have the local definition

$$I_{\alpha\beta}Z^\alpha_1 Z^\beta_2 = \varepsilon^{A'B'}\pi_{1A'}\pi_{2B'} = \pi_{1}{}^{0'}\pi_{2}{}^{1'} - \pi_{1}{}^{1'}\pi_{2}{}^{0'}$$

so we might expect to average this expression over $\mathcal{S}$ in a general $\mathcal{M}$ (the expression being certainly not constant in general), together with some possible involvement of the η-factor. The resulting "averaged" $I_{\alpha\beta}$ will not generally be of the Minkowski kind (i.e. "simple": $\det(I_{\alpha\beta}) = 0$) but will be a deSitter-type infinity twistor. In the particular situations that arise when $\mathcal{S}$ is taken to either spacelike or null infinity, we do ineed get a Minkowski-type infinity twistor, and the previously mentioned Hermiticity condition for $A_{\alpha\beta}$ indeed holds. However, the general problem of a Hermiticity relation for $A_{\alpha\beta}$ with a general (even uncontorted) $\mathcal{S}$ remains unsolved.

We can see, from various considerations, that if there is such a Hermiticity relation, it certainly has to be generally with a deSitter-type infinity twistor and not a Minkowski-type one. One such consideration arises from an examination of Tod's (1983) results for $\mathcal{S}$ lying in a conformally flat hypersurface of time-symmetry. Here we find that when $\mathcal{S}$

surrounds a source region $\mathcal{R}_i$ (once) we have the same angular momentum twistor $A_{i\alpha\beta}$, irrespective of the precise location of $\mathcal{S}$, and similarly if it surrounds $\mathcal{R}_i \cup \mathcal{R}_j$ (once) we always get

$$A_{ij\alpha\beta} = A_{i\alpha\beta} + A_{j\alpha\beta} \quad .$$

However, defining the corresponding rest-masses via the equation

$$m^2 = -\frac{1}{2} A_{\alpha\beta}\bar{A}^{\alpha\beta}$$

(which is well-defined since $\mathcal{S}$ is uncontorted), we find

$$m_{ij} < m_i + m_j \quad .$$

Such an inequality would be impossible if the various angular momentum twistors always had the form

$$A_{\alpha\beta} = \begin{pmatrix} 0 & p_{A'}{}^{B'} \\ p^{A'}{}_{B'} & 2i\mu^{A'B'} \end{pmatrix}$$

(p_a being the 4-momentum and $\mu^{A'B'}$ describing the angular momentum) which occurs when the Hermiticity relation for $A_{\alpha\beta}$ holds with a standard Minkowski-type infinity twistor. However an entry other than zero can occur in the upper left-hand corder with a deSitter-type Hermiticity relation, and this is indeed what happens here.

An even more elementary illustration occurs with the Schwarzschild space-time, where we consider surfaces $\mathcal{S}$ sharing the spherical symmetry of the space-time and lying in the t = 0 hypersurface $\mathcal{H}$. This hypersurface is one which admits Tod's (1984) 3-surface twistors, which implies that (as in the more general time-symmetric case just considered) the spaces of 2-surface twistors for any $\mathcal{S}$ lying in $\mathcal{H}$ can be canonically identified with one another - and with the space $\mathbb{T}(\mathcal{H})$ of 3-surface twistors for $\mathcal{H}$. However, there is no appropriately unique infinity twistor for

$\mathcal{H}$. If we start with $\mathcal{I}$ out at spacelike infinity i^0, we obtain an infinity twistor which represents that point: i^0. But if we move $\mathcal{I}$ continuously (and spherically symmetrically) inwards from i^0 until it reaches the horizon, and then out again on the opposite Kruskal sheet, we finally reach the "opposite" spacelike infinity point $\hat{i}^0$. Now from the point of view of $\mathbb{T}(\mathcal{H})$ (i.e. of the conformal geometry of $\mathcal{H}$) the point $\hat{i}^0$ seems like the origin when "viewed" from i^0, whereas i^0 itself seems like the origin when "viewed" from $\hat{i}^0$. Thus we get two distinct Minkowski-type (i.e. simple) corresponding infinity twistors, say

$$I_{\alpha\beta} \quad \text{and} \quad \hat{I}_{\alpha\beta} \qquad .$$

For an arbitrary (spherically symmetrical) location of $\mathcal{I}$, we must, by symmetry get some linear combination of these two. This combination passes smoothly from $I_{\alpha\beta}$ itself to $\hat{I}_{\alpha\beta}$ itself and, in between, we must (and do) get infinity twistors which are of deSitter type.

In fact the Hermiticity relation for $A_{\alpha\beta}$ does hold with this deSitter-type infinity twistor in this case, but it is a very special situation and we do not know what to expect generally. However, the consideration of deSitter-type infinity twistors seems inescapable. Accordingly there will, in general, be no way of projecting out the 4-momentum from the full complex of quantities in $A_{\alpha\beta}$.

References

Jeffryes, B.P. (1984) Two-surface twistor and conformal embedding in Asymptotic Behaviour of Mass and Spacetime Geometry, Proc. Corvalis Oregon 1983, ed. F.J. Flaherty (Springer-Verlag, Berlin)

Penrose, R. (1982) Quasi-local mass and angular momentum in general relativity, Proc. Roy. Soc. London A381, 53-63

Penrose, R. & Rindler, W. (1986) Spinors and Space-Time, Vol. 1, Two-Spinor Calculus and Relativistic Fields (Cambridge University Press, Cambridge)

Penrose, R. & Rindler, W. (1986) Spinors and Space-Time, Vol. 2, Spinor and Twistor Methods in Space-time Geometry (Cambridge University Press, Cambridge)

Shaw, W.T. (1983) Twistor theory and the energy-momentum and angular momentum of the gravitational field at spatial infinity, Proc. Roy. Soc. London A390, 191-215

Tod, K.P. (1983) Some examples of Penrose's quasi-local mass construction, Proc. Roy. Soc. London A388, 457-77

Tod, K.P. (1984) Three-surface twistors and conformal embedding, Gen. Rel. Grav. 16, 435-43

181
FUNDAMENTAL ASYMMETRY IN PHYSICAL LAWS

The mathematical heritage of Hermann Weyl (1987), 317–28. *Proc. Sympos. Pure Math.*, **48**. American Mathematical Society, Providence, 1988

This was the second meeting I attended in honour of Hermann Weyl. See also Chapter 169 and the footnote on p. [323] of the present article.

 Fundamental asymmetry in physical laws. *The mathematical heritage of Hermann Weyl* (1987), 317–28. *Proc. Sympos. Pure Math.*, **48**. 1988

Proceedings of Symposia in Pure Mathematics
Volume 48 (1988)

Fundamental Asymmetry in Physical Laws

R. PENROSE

Weyl and asymmetry in physics. The concept of *symmetry* was clearly a central theme in much of Hermann Weyl's work; and, as is well known, he wrote a beautiful semi-popular book, with that title, late in his life (Weyl 1952). It would be clear from this book alone that, along with being a great mathematician, Weyl was a philosopher, and he much appreciated the arts. But in addition to these other interests, he was a physicist of definite distinction, having had deep insights into the workings of the natural world and having introduced several important and highly influential ideas (e.g., in particular, gauge theory). Symmetry was an underlying theme in many of Weyl's contributions to physics. Indeed, one of Weyl's most significant early physical realizations was the importance of symmetry in quantum theory, and his book *The theory of groups and quantum mechanics* (Weyl 1931) was an important landmark.

Perhaps almost as significant for physics as his recognition of the role of symmetry was his suspicion of a possible fundamental role for asymmetry in the laws governing physical behaviour. Weyl seems to have been particularly intrigued by the *discrete* symmetries of space-reflection, particle-antiparticle interchange, and time-reversal—these being the now well-known operations denoted by **P**, **C** and **T**, respectively. He pointed out (Weyl 1931) that Dirac's "hole" theory implied that the holes must have the *same* mass as the electrons—and so could not be protons, as Dirac had initially hoped (invariance under **C**). He considered the question of invariance under **P**, **C** and **T**, noting that such invariance was an additional issue to the normal Lorentz invariance requirements. In 1929 Weyl introduced the wave equation that now bears his name,[1]

$$(\partial/\partial t-\partial/\partial z)\nu_0=(\partial/\partial x-i\partial/\partial y)\nu_1, \qquad (\partial/\partial t+\partial/\partial z)\nu_1=(\partial/\partial x+i\partial/\partial y)\nu_0,$$

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 83C45, 81B05.

[1]Ironically, it is the physicists who pay tribute to the mathematician Weyl by referring to this (massless) equation as the "Weyl equation", whereas the mathematicians honour a physicist, calling it the "Dirac equation"! (Strictly, the mathematicians are historically more correct, since Dirac had hit upon this wave equation first—and discarded it for what were good physical reasons at the time—before introducing his more famous *massive* Dirac equation; cf. Dirac 1928, 1982.

©1988 American Mathematical Society
0082-0717/88 $1.00 + $.25 per page

and which is now the standard equation used to describe the (massless) neutrino. The lack of space-reflection invariance of this equation had got him into trouble with the established physicists, particularly Pauli, since physical laws had been thought, at that time, to be completely reflection-invariant. The fact that **P** is actually *not* a symmetry of *weak* interactions (and nor are **C**, **CT**, **PT**) was not discovered until 1957 (Wu et al. 1957, cf. Lee and Yang 1956), about a year after Weyl's death.

Almost symmetries in nature. It is a puzzling fact that there are basic symmetries in nature that are only approximately true—or true only for some interactions but not for others. The **P** and **C** symmetries are examples of this, being symmetries of the strong, electromagnetic and gravitational interactions, but not of the weak interactions. There is also a *single* phenomenon known, for which **CP** is not an exact symmetry—and neither is **T**: the decay of the K^0 particle (Christenson et al. 1964)! Such puzzling "almost symmetry" in nature is a feature of other symmetry operations, also. In the modern theories of particle classification, symmetry groups play a fundamental role. The original example of such a group was SU(2), which referred to a concept called *isotopic* (or, more correctly, *isobaric*) *spin*, introduced in the 1930s. Elements of this SU(2) group can "rotate" protons into neutrons. Strong interactions respect this symmetry, whereas weak and electromagnetic interactions do not. This SU(2) group was extended, in 1962, to SU(3) by Gell-Mann, Ne'eman and others (cf. Gell-Mann and Ne'eman 1964). Again the group is (essentially) respected by strong, but not weak or electromagnetic, interactions. Nowadays, theoreticians often use much larger symmetry groups. The idea is now to have a more unified type of theory in which strong, weak and electromagnetic interactions are *all* to be thought of as different aspects of *one* grand scheme. Thus, according to the various "grand unified theories" (abbreviated GUT), all these different interactions should come together as different aspects of a single concept, and transform among themselves under the action of the proposed group. As yet there is no consensus as to which GUT theory would be most appropriate, and none of these theories appears to have significant experimental support so far. However, the more modest *electroweak* theory (using the group $SU(2) \times U(1)$), due to Salam, Ward, Weinberg and Glashow, which just unifies weak interactions with electromagnetic ones, has indeed had some impressive successes, and the theory is now regarded as standard.

Spontaneous symmetry breaking. In all of these unified theories, the entire symmetry group of the theory is *not* an exact symmetry of the relevant observed physical phenomena. A crucial idea is needed, in order to make such a theory work. This is to invoke the procedure of *spontaneous symmetry breaking*. Accordingly, when the ambient temperature is greater than a certain critical value, the physical phenomena do indeed respect the entire symmetry; but below this critical temperature, symmetry breaking sets in, and only some smaller symmetry group remains, and is respected by the observed phenomena. An example

 Fundamental asymmetry in physical laws. *The mathematical heritage of Hermann Weyl* (1987), 317–28. *Proc. Sympos. Pure Math.*, **48**. 1988

taken from standard everyday physics is often cited as a particular instance of this type of phenomenon: the example of *ferromagnetism*. Above the critical temperature, a piece of iron will not spontaneously exhibit magnetic properties. Though the various atoms behave individually like little magnets, and though the atoms have a tendency to align themselves with the magnets in the neighbouring atoms, the random motions induced by the ambient temperature are too great to allow the atoms to align themselves collectively. But when the temperature drops below this value, these random motions are not sufficient to overcome the tendency for alignment, and a lower energy configuration is achieved if the atoms *are* all aligned with one another. Now imagine a solid spherical ball of this iron which is initially above the critical temperature, but is then allowed to cool. As the iron cools below the critical temperature, it becomes energetically favoured that the ball should, as a whole, become a magnet, pointing in some particular direction. What direction does it choose? Assuming that there is no background magnetic field, the direction must be chosen *at random*. The original rotational symmetry is now broken. The physics as a whole does have rotational symmetry, and above the critical temperature the equilibrium state is indeed spherically symmetric. But below this temperature, the equilibrium state does *not* have full rotational symmetry, but only axial symmetry. The symmetry has been *broken*—from SO(3) down to SO(2).

More abstractly, situations of this general kind are frequently described in terms of a "Mexican hat potential", e.g.

$$V = V_0 + (1 - r^2)^2,$$

where r is the distance from the centre in, say, a Euclidean k-space representing the different possible states of a system, depicted horizontally, and the potential energy V is depicted as the height. The lowest energy is achieved around the "rim of the hat", namely the S^{k-1} given by $r = 1$, this energy being *degenerate*. At low temperature, the system settles into this degenerate energy, but no one point on the S^{k-1} is preferred over any other. Which point is actually chosen is a matter of pure chance, and when the choice is made, the symmetry of the system is reduced to that of S^{k-2}, providing us with a typical example of spontaneous symmetry breaking.

This kind of procedure is a very basic ingredient, common to almost all modern unified theories. Consider electroweak theory. The view would be that in the very early stages of the universe when, according to the standard "hot big bang" theory, it was vastly smaller and at a far higher temperature than it is today, the forces of electromagnetism and weak interactions were on an equal footing with each other as part of one comprehensive symmetrical scheme. As the universe expanded, it would have cooled and dropped below the critical temperature. At that stage, a "direction" would be singled out in the space on which the symmetry acts, and the electromagnetic interaction would become distinguished from weak interactions—i.e., photons distinguished from W-bosons, the latter

acquiring masses—and the SU(2) × U(1) symmetry group of the theory would become reduced to (a diagonal) U(1).

Two possible viewpoints. One may contrast this idea of a symmetry-breaking process with a completely opposed point of view, according to which higher symmetry is not taken as fundamental at all, but just an approximation, the underlying exact theory being not supposed actually to possess this higher symmetry. Approximate symmetries of this kind are not unfamiliar in situations in which a good underlying theory happens to be known. An example is given by the theory of atomic structure, where groups and group representations can play important roles, even when the group involved provides only an approximate symmetry of the entire physics (cf. Dyson 1966). In this example, the underlying theory comes from standard quantum mechanics and electromagnetism. In such a situation, one does *not* expect the approximate theory to become exact when the temperature exceeds a critical value, so the implications of such a viewpoint are quite different. For some reason, viewpoints of this kind do not appear to find much favour with modern particle physicists, and the alternative of symmetry-breaking is the viewpoint almost universally adopted. (One of the reasons for this has to do with renormalization theory. A good way of producing a quantum field theory for which troublesome infinities cancel one another out is to have a gauge theory with a high degree of symmetry, the lack of such symmetry at the *observed* level being attributed to a symmetry-breaking phenomenon like that discussed above—referred to as a *Higgs mechanism* 1966).

We thus have two basic alternative viewpoints with regard to the deviations from perfect symmetry in physics:

S: spontaneous symmetry-breaking,

N: underlying physics does not have the full symmetry.

Of course, for some physical phenomena, S can be more appropriate, while for others it can be N. I do not mean to imply that either one of these mechanisms is universally to be preferred in physics generally—and we have seen above that there are examples in ordinary physics where each of these holds. However, the issue seems to become more ideological, if one believes that one is approaching a fundamental truth underlying physical laws. Does one believe in an underlying physics in which some large Lie group (or perhaps some large finite group—such as the "monster group", as has occasionally been suggested) is a fundamental symmetry of the scheme? Or is a large symmetry group, just in itself, something of a complication? Might one prefer, instead, some underlying scheme in which symmetry, as such, has no particular fundamental role to play? My own personal preference is indeed for the latter possibility—i.e., for N rather than S, at the fundamental level. I do not see why symmetry, as such, should be fundamentally important, though many basic mathematical structures do indeed possess important symmetry properties. Symmetry is very important for the description and understanding of such basic structures, but I do not see why it should be put in there at the start. This seems to be a minority view among present-day particle

physicists—for symmetry does indeed have a special beauty of its own—but I would like to feel that Weyl might actually have been on my side!

The discrete symmetries P, C and T. Let us examine the symmetries **P**, **C** and **T**, in relation to the two possible viewpoints **S** and **N**. (Here it is not quite so unfashionable to adopt **N**, at the fundamental level, as it is with the continuous symmetries.) It should be remarked, first, that there is an important theorem, known as the **PCT**-theorem, which asserts that the product **PCT** of all three symmetries **P**, **C** and **T** should be an exact symmetry for any quantum field theory subject to certain "physically reasonable" assumptions. These assumptions have to do with causality and positivity of energy, and invariance under the nonreflective transformations of the Poincaré group (which is the symmetry group of Minkowski space-time). Note the necessity of this last assumption. In my view this could be significant, because this assumption is explicitly violated by general relativity. We do not yet know the correct way to combine the basic ideas of quantum field theory with the principles of curved-space geometry—principles which are fundamental to Einstein's theory. I shall be indicating shortly my reasons for believing that, in fact, **PCT** ought to be *false* for the correct quantum gravity theory!

I have remarked that, along with **CP**, the symmetry **T** is false for K^0-decay. In fact the operation of **T** is a puzzle for physics in other ways. Let us disregard **T**-violation in K^0-decay for the moment. That is a very tiny effect, and not directly significant for any known action shaping the universe. We now find that there is a striking discrepancy between the symmetry principles which hold for the large-scale behaviour of macroscopic objects and those which seem to hold at the level of submicroscopic physics and exact physical laws. Invariance under **T** holds for the latter, whereas time-asymmetry is manifest for many large-scale phenomena.

Entropy and the second law. The most far-reaching of these time-asymmetries is that provided by the *second law of thermodynamics*. According to the second law, the quantity known as *entropy*, which measures the manifest disorder of a system, increases with time. (The entropy of a state of a system can be taken to be the logarithm of the phase-space volume of all the states which are to be regarded as macroscopically indistinguishable from the given state. This seems somewhat subjective, but this subjectiveness is not the important issue for the points that I wish to make. Reasonable changes in one's viewpoint as to which states are to be regarded as "macroscopically indistinguishable" make virtually no difference to the entropy values in most circumstances.) If a system is given to us in some state having a low value for its entropy, then it is not really surprising that its most probable behaviour in the future will be such as to increase this entropy. (For the point in phase space which represents the system will be likely to move into larger and larger phase-space regions as time progresses.) However, the *puzzle* that the second law presents us with is that if we take the same given low-entropy state and ask for its most likely behaviour in the *past* direction in

time, then—in the actuality of our universe—the system gets into more and more manifestly improbable states as time progresses farther and farther back in time. In the absence of some other given constraint on the system in the remote past, we ought to expect that the most likely way to achieve some given low-entropy state would be to approach it with other *more* probable (i.e., higher entropy) states first, and achieve the desired given low-entropy configuration as a local minimum. This is blatantly *not* what happens in our universe. The entropy of the whole universe decreases and decreases in the past direction in time, until it reaches its minimum value at the very beginning of time. (If our universe is infinite, then one needs to be a little careful about talking about its "total entropy". But again, the technical problems concerned with precise definitions are not the essential issue here.)

The "beginning of time" occurred with the big bang. So the constraint on the universe which set the second law on its course must have occurred then. One point should be made clear. This constraint was not simply that the universe started out as very "tiny". There were as many degrees of freedom available at the beginning of time as there are at any other time. The entropy did not *need* to be low simply because the universe was small. Perhaps the easiest way to see this is to consider a universe model which is collapsing. The entropy can continue to go up, but one finds that most of the entropy resides in black holes. This black-hole entropy can be calculated by use of a formula due to Bekenstein (1974) and Hawking (1978) which asserts that (in natural units: the velocity of light, Planck's constant/2π, Newton's gravitational constant and Boltzmann's constant are all set equal to unity) this entropy is just $A/4$, where A is the surface area of the hole's horizon. In the final collapse, these black holes all come together, and the space-time singularities which lie at their cores (where curvatures diverge to infinity) all congeal into one final singularity—referred to as the big crunch. The entropy in a big crunch singularity would be far larger than that in the big bang. Essentially, this entropy provides a measure of the complication of the singularity. The big bang singularity was indeed highly constrained—and one can more or less understand the nature of this constraint, namely that the Weyl conformal curvature was zero—or at least very small in comparison with the remaining Ricci curvature (Penrose 1979a). Such a constraint cannot apply to the big crunch, or to the singularities inside black holes generally. If such a constraint did apply, then one would have severe violations of the second law of thermodynamics.

The difference between these latter types of singularity and the singularity of the big bang is that they are *future* singularities (i.e., timelike curves can enter them in the future direction) whereas the big bang was a *past* singularity (timelike curves can come out of it in the future direction). The rules for past singularities must be different from those which apply at future singularities. Past singularities (or, rather, the only past singularity that we know about, namely the big bang) seem to be highly constrained in having zero (or else very

small) Weyl curvature. I refer to this constraint as the *Weyl curvature hypothesis* (cf. Penrose 1981, 1986a, b).[2]

The common belief is that it should be *quantum gravity* which determines whatever these rules are—and so it should presumably be quantum gravity which provides the Weyl curvature hypothesis (and hence the second law). However, we note that the Weyl curvature hypothesis is not invariant under **T** (nor under **PT**, nor **CT**, nor even under **PCT**). This seems to imply that quantum gravity theory is not invariant under these symmetries either. At least this would be the view consistent with **N**. Some people might prefer an explanation more in line with **S**. However, it is difficult to see how the ideas of symmetry breaking can apply here. Symmetry breaking, in the ordinary sense, refers to a system "settling down" into an asymmetrical configuration in the normal future direction in time—and it thereby *makes use of* the second law of thermodynamics rather than being able to provide an *explanation* for that law! It seems that the asymmetry in time is already being assumed in the standard symmetry-breaking mechanisms, and I do not see how **S** can be invoked to explain time-asymmetry.

Inflationary universes. I should make some comment concerning the fashionable "inflationary models" (cf. Hawking, Gibbons and Siklos 1986) which purport to explain some of the puzzling features of the early universe in terms of mechanisms which depend heavily upon an **S**-type process. The idea is that in the extremely early stages of the big bang, the universe expanded by an absolutely enormous factor because of some phase change in the vacuum at very early times, using some **S** mechanism. This enormous expansion is supposed to explain the big bang's spatial uniformity of expansion, which is estimated to be "fine-tuned" to a factor of at least one part in 10^{60}. However, inflation provides no real explanation, since the physics in the inflationary models is taken to be consistent with **T**. Thus, the inflationary scheme (in reverse) should apply equally to a generic collapsing universe—which it blatantly need not, because of the black-hole singularities. A time-asymmetric constraint is still needed at space-time singularities, and we appear to be back with the Weyl curvature hypothesis, as before, for the initial singularity. An explanation for the time-asymmetric second law is *not* provided by the inflationary models. In fact the puzzle of fine tuning to one part in 10^{60} is totally dwarfed by the constraint needed to remove the time-reverses of black holes. For a closed universe with about 10^{80} baryons such a constraint would need fine tuning to about one part in 10^{123}. For a larger universe, this fine tuning would need to be even more precise. It is this that we need the Weyl curvature hypothesis for, whether or not an inflationary phase of the universe actually occurred.

[2]Weyl himself seems to have been aware that some important constraint was needed for the early universe. According to "Weyl's postulate" (cf. Bondi 1952) the idealized substratum particles must have world-lines which are orthogonal to spacelike hypersurfaces. This implies some degree of uniformity for the big-bang singularity. There seems to be some loose connection between Weyl's postulate and what I am calling the Weyl curvature hypothesis.

It should be remarked that the inflationary models are also invoked in order to explain other seeming puzzles, namely the absence of a proliferation of such extravagant phenomena as "domain walls", "cosmic strings", "axions" and "monopoles". These are all animals of the menagerie conjured up by a belief in the reality of **S**, in current unified gauge theories. If we eventually find that we do not need **S**, then the inflationary models will not be needed either!

T-noninvariance at the fundamental level? I have indicated by belief that a fundamental theory of physics ought to be more in line with **N** than with **S**. Can one provide a basic theory of physics which, at the *fundamental* level, is *not* invariant under **P**, **C**, **T**, or under any nontrivial combination of these three? In my opinion this will be ultimately necessary. I have indicated that I believe that the correct theory of quantum gravity, when it is found, should be time-asymmetric. I believe, also (cf. Penrose 1981, 1986a, 1987), that such a theory ought to give some kind of objective picture of what happens when "collapse of the wave function" takes place in ordinary quantum mechanics. I have argued elsewhere that this procedure also must be time-asymmetric (Penrose 1987). According to the standard quantum-mechanical prescription, quantum amplitudes are converted into probabilities—by our taking the squares of their moduli—whenever an "observation" (and hence "wave-function collapse") is deemed to have taken place. This procedure works marvellously, and is fully consistent with observational facts, provided that it is employed in the normal future direction. However, it would give completely the wrong answers if it were to be applied in the past direction!

The procedures of unitary evolution (i.e., Schrödinger equation, or equivalent), on the other hand, are completely time symmetric. Classical general relativity, also, is an entirely time-symmetric theory. It would seem, therefore, that in the absence of any mechanism of the **S** type, quantum gravity ought also to be time-symmetric if it is simply some kind of union between these two time-symmetric procedures. It is thus my opinion that the union between quantum theory and gravitational theory will entail an actual change in the structure of quantum theory, where the unified theory will be time-asymmetric at the fundamental level—in accordance with **N**—and some objective and mathematically well-defined procedure will take over the rather unsatisfactory "collapse of the wave function" that is a feature of present quantum theory.

P- and C-noninvariance: twistor theory. What about the symmetries **P** and **C**? Can asymmetry with respect to these operations also be built in at the fundamental level? Some modern string theories attempt to do just this (e.g., the proposal of so-called "heterotic strings", though I have found some difficulty, myself, in understanding just how these theories are to be interpreted. A radically different approach from string theory, albeit with certain common ideas (cf. Shaw 1986, Hughston and Shaw 1987a, 1987b), is that provided by *twistor* theory (see Huggett and Tod 1985, Penrose and Rindler 1986, Ward

and Wells 1988). According to the twistor approach, the structures of holomorphic geometry are taken as fundamental. However, the basic geometry is not that of space-time, but of twistor space (or of portions of products of twistor spaces). At the most visualizable level, twistor space can be regarded as describing, in a holomorphic way, the space of (straight) light rays in four-dimensional Minkowski space-time **M**. The light rays constitute a five-dimensional system, but the 5-manifold describing them imbeds naturally as a real hypersurface **PN**, in a complex projective 3-space **PT**. Here, **PT** is the *projective* version of the full twistor space **T**, which is a complex vector space with a pseudo-Hermitian structure of signature $(++--)$, the null elements of which determine **PN**. The points of **M** are interpreted as complex projective lines, in **PT**, which lie in **PN**.

There is some immediate relation to string theory in four-dimensional space-time (at least for "string histories" whose induced metric from the space-time is definite rather than Lorentzian), since it turns out that holomorphic curves in **PT** describe general solutions of the "minimal surface equation" in **M**. (This construction has been studied by Shaw 1986 and generalized to other induced-metric signatures and to other space-time dimensions—up to ten—by Hughston and Shaw 1987a, 1987b.) However the underlying philosophy of twistor theory is different from that of string theory. With twistor theory one tries to build up the space-time points themselves as secondary structures, the twistor space being regarded as more fundamental. In standard twistor theory, the space-time is four-dimensional and the signature of the space-time metric is Lorentzian, though many higher-dimensional generalizations and departures from this original scheme have been studied.

One of the underlying thoughts of twistor theory is that symmetry groups are not necessarily to be taken as fundamental, but they may arise as "incidental" structures. The basic example of this sort of thing is provided by the complex projective lines in **PN** which represent points of **M**. The *points* of such a line constitute a Riemann sphere in **PN** representing the light rays in **M** passing through a particular point **x** in **M**. Thus, this Riemann sphere represents the "sky" or celestial sphere of an observer situated at **x**. Now, this celestial sphere has a natural structure as a one-dimensional complex manifold (i.e., it is naturally an oriented conformal real 2-manifold). This structure is preserved as we pass from one observer at **x** to another. The group which preserves this structure is $\mathrm{PL}(2,\mathbf{C})$, which is the same as the *restricted* (i.e., nonreflective) *Lorentz group*. This group may be regarded as the most important symmetry group of physics, and it is seen as arising here as the simplest symmetry group of a "structureless" complex manifold: the group of symmetries of a complex 1-manifold of simplest (i.e., S^2) topology. This kind of idea also extends (though perhaps less impressively) to certain higher-dimensional symmetry groups.

Massless fields and cohomology. Linear massless space-time fields—or, rather more appropriately, wave functions—can be represented very neatly in terms of *sheaf cohomology* elements for appropriate regions of **T** (or of **PT**).

There is a relation between the homogeneity degree n of the functions—called *twistor functions*—providing the required representative (Čech) cocycle—and the *helicity* s of the field:

$$n = -2s - 2$$

(where units are chosen for which Planck's constant/2π is taken as unity). This means, in particular, that the *right-handed photon* (described by a wave function which is a self-dual Maxwell field) has a twistor function of homogeneity degree $n = -4$, whereas a *left-handed photon* (anti-self-dual Maxwell field) has a twistor function of degree $n = 0$. For a Weyl *neutrino* (left-handed) we have $n = -1$, while for an *anti-neutrino* (right-handed), $n = -3$. If we take gravitons to be described by linearized solutions of Einstein's vacuum equations (i.e., vanishing Ricci tensor), then we find that the *right-handed graviton* (wave function describing a self-dual linearized Weyl tensor) has a twistor function of homogeneity degree $n = -6$, while a *left-handed graviton* (anti-self-dual linearized Weyl tensor) has a twistor function of degree $n = +2$.

Note that there is a fundamental left-right asymmetry in this description. A space-reflection (operation of P) would merely reverse the sign of the helicity,[3] but this does something more complicated to the homogeneity degree of the twistor function, namely effect the interchange

$$n \longleftrightarrow -n - 4.$$

This is the kind of thing that can make a significant difference in a cohomology description, though such a difference does not actually show up for the (twistor) cohomology group elements describing free wave functions.

Left-right asymmetry in nonlinear twistor descriptions. A more serious left-right asymmetry arises with the *nonlinear* versions of these cohomology elements. In close relation to the Kodaira-Spencer (1958) theory of deformations of complex manifolds and the theory of holomorphic vector bundles, such cohomology elements can be "exponentiated", in certain cases, to give nonlinear fields (or "nonlinear wave functions"). In the *nonlinear graviton construction* (Penrose 1976) this is achieved for the case $n = 2$, and provides the general (complex) solution of the Einstein vacuum equations with anti-self-dual curvature, while in the *Ward* (1977) *construction* (cf. also Atiyah and Ward 1977, Atiyah, Hitchin, Manin and Drinfeld 1978), the case $n = 0$ is generalized to provide the general solution of the anti-self-dual Yang-Mills equations. For a more complete physical theory, one would need to be able to code *self*-dual nonlinear fields into deformations of twistor structures in addition to anti-self-dual ones (this is the *googly* problem, cf. Penrose 1979b) and then to combine the two to give the full description of a nonlinear field which need be neither self-dual nor anti-self-dual.

The self-dual nonlinear fields seem to be very much more difficult to handle in terms of the standard twistor space (the googly problem) than do the

[3]The different roles of P and C are a little difficult to disentangle in all this. There is actually a choice, in the twistor description which, via the sign of the given Hermitian form, serves to single out one half of PT\PN, denoted by PN^+, as preferred over the other.

anti-self-dual ones. A description in the gravitational case seems to be possible, but it involves some rather strange ideas, involving the "exponentiation" of *relative* (second) cohomology elements (Penrose 1986b) and leading to peculiar nonlinear deformations up the fibres of $\mathbf{T} \to \mathbf{PT}$, providing a type of locally extendible "conical singularity". There is a blatant left-right asymmetry in this that many would no doubt find disturbing—since Einstein's vacuum equations, for example, are perfectly left-right symmetric. The asymmetry arises because of a fundamental choice having been made right at the start: *twistor space has been chosen in preference to dual twistor space.* Had we started with dual twistor space instead, the right-handed (self-dual) cases would be the easy ones and the left-handed (anti-self-dual) cases, the difficult ones. An alternative route to follow is that provided by the *ambitwistor* descriptions, which work with a space representing the complex null geodesics in complexified space-time which can, in a certain sense, be realised as a subspace of the product space of twistor space with its dual. This route has been pioneered by Isenberg, Yasskin and Green (1978), LeBrun (1983), Witten (1978, 1986) and Manin (1982). It does not exhibit left-right asymmetry, and has its own disadvantages as well as advantages. Its connection with twistor wave functions is less clear than is the case for the original asymmetrical twistor approach.

In my own view, despite its technical difficulties, the fundamentally asymmetrical approach is likely to be ultimately the most fruitful. Only time—and a great deal of hard work, insight and good luck—will tell!

REFERENCES

M. F. Atiyah and R. S. Ward (1977), *Instantons and algebraic geometry*, Comm. Math. Phys. **55**, 111–124.

M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Yu. I. Manin (1978), *Construction of instantons*, Phys. Lett. **65A**, 185–187.

J. D. Bekenstein (1974), *Generalized second law of thermodynamics in black-hole physics*, Phys. Rev. **D9**, 3292–3300.

H. Bondi (1952), *Cosmology*, Cambridge Univ. Press, Cambridge.

J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay (1964), *Evidence for the* 2π *decay of the* K^0 *meson*, Phys. Rev. Lett. **13**, 138–140.

P. A. M. Dirac (1928), *The quantum theory of the electron*, Proc. Roy. Soc. London **A117**, 610–624; part II, ibid. **A118**, 351–361.

—— (1982), *Pretty mathematics*, Int. J. Theor. Phys. **21**, 603–605.

F. J. Dyson (1966), *Symmetry groups in nuclear and particle physics*, Benjamin, New York.

M. Gell-Mann and Y. Ne'eman (1964), *The eightfold way*, Benjamin, New York/Amsterdam.

S. W. Hawking (1978), *Particle creation by black holes*, Commun. Math. Phys. **43**, 199–220.

S. W. Hawking, G. Gibbons, and S. Siklos (1986), *The very early universe*, Cambridge Univ. Press, Cambridge.

P. Higgs (1966), *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145**, 1156–1163.

L. P. Hughston and W. Shaw (1987a), Classical Quantum Gravity **4**, 869.

—— (1987b), *Classical strings in ten dimensions*, Proc. Roy. Soc. London **A414**, 423–431.

J. Isenberg, P. B. Yasskin, and P. S. Green (1978), *Non-self-dual gauge fields*, Phys. Lett. **78B**, 462–464.

K. Kodaira and D. C. Spencer (1958), *On deformations of complex analytic structures. I, II*, Ann. Math. **67**, 328–401, 403–466.

S. A. Huggett and K. P. Tod (1985), *An introduction to twistor theory*, London Math. Soc. Student Texts (L.M.S. publ.).

T. D. Lee and C. N. Yang (1956), *Question of parity conservation in weak interactions*, Phys. Rev. **104**, 254–258.

C. R. LeBrun (1983), *Spaces of complex null geodesics in complex Riemannian geometry*, Trans. Amer. Math. Soc. **284**, 601–616.

Yu. I. Manin (1982), *Gauge field and cohomology of analytic sheaves*, Twistor Geometry and Non-Linear Systems (H. D. Doebner and T. D. Palev, eds.), Springer-Verlag, Berlin.

R. Penrose (1976), *Non-linear gravitons and curved twistor theory*, Gen. Rel. Grav. **7**, 31–52.

—— (1979a), *Singularities and time-asymmetry*, General Relativity: An Einstein Centenary (S. W. Hawking and W. Israel, eds.), Cambridge Univ. Press, Cambridge.

—— (1979b), *A googly graviton?* Advances in Twistor Theory (L. P. Hughston and R. S. Ward, eds.), Pitman Press, Boston, pp. 168–176.

—— (1981), *Time-asymmetry and quantum gravity*, Quantum Gravity 2 (C. J. Isham, R. Penrose and D. W. Sciama, eds.), Oxford Univ. Press, Oxford.

—— (1986a), *Gravity and state-vector reduction*, Quantum Concepts in Space and Time (R. Penrose and C. J. Isham, eds.), Oxford Univ. Press, Oxford.

—— (1986b), *Hermann Weyl, space-time and conformal geometry*, Hermann Weyl 1885–1985 (K. Chandrasekharan, ed.), I. T. H., Zurich and Springer-Verlag, Berlin.

—— (1987), *Newton, quantum theory and reality*, 300 Years of Gravity (S. W. Hawking and W. Israel, eds.), Cambridge Univ. Press, Cambridge.

R. Penrose and W. Rindler (1986), *Spinors and space-time*. Vol. 2: *Spinor and twistor methods in space-time geometry*, Cambridge Univ. Press, Cambridge.

W. T. Shaw (1986), Classical Quantum Gravity **3**, 753–761.

R. S. Ward (1977), *On self-dual gauge fields*, Phys. Lett. **61A**, 81–82.

R. S. Ward and R. O. Wells, Jr. (1988), *Twistor geometry and field theory*, Cambridge Univ. Press, Cambridge.

H. Weyl (1918), *Gravitation und Electrizität*, Sitz. Ber. Preuss. Ak. Wiss., 465–480.

—— (1929), *Elektron und Gravitation*. I, Z. Phys. **56**, 330–352.

—— (1931), *The theory of groups and quantum mechanics*, Methuen, London.

—— (1952), *Symmetry*, Princeton Univ. Press, Princeton, N. J.

E. Witten (1978), *An interpretation of classical Yang-Mills theory*, Phys. Lett. **77B**, 394–398.

—— (1986), Nuclear Phys. **B 266**, 245–264.

C. S. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson (1957), *Experimental test of parity conservation in beta decay*, Phys. Rev. **105**, 1413–1415; *Further experiments on beta decay of polarized nuclei*, Phys. Rev. **106**, 1361–1363.

MATHEMATICAL INSTITUTE, OXFORD, UNITED KINGDOM

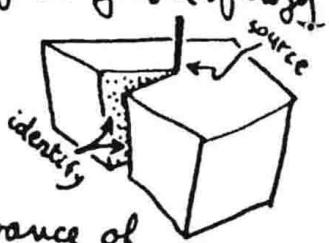
182
TOPOLOGICAL QFT AND TWISTORS: HOLOMORPHIC LINKING

Twistor Newsletter **27** (1988), 1–4

I feel that there is something in these ideas that needs to be explored much more fully.

Topological QFT and Twistors: Holomorphic Linking

There is much current interest in the idea of "topological quantum field theories" (TQFT) particularly in relation to pure mathematical applications: Fleur cohomology, Donaldson invariants, Jones polynomials, etc. However the physical relevance of such "field theories" is more problematical. The salient feature of a TQFT is that there are no field equations — the dynamics vanishes altogether locally, and the theory can be formulated to be independent of the choice of metric. This is very attractive from the mathematical point of view – where one may be interested in global (differential) topological properties rather than properties depending upon a chosen metric or (physical) field. But physical relevance remains unclear. A noteworthy example, where field equations vanish locally, is provided by (2+1)-dimensional "gravity": in "vacuum" we have $R_{ab}=0$ whereas the Weyl curvature necessarily vanishes ($C_{abcd}=0$) in 3-dimensions [cf. E. Witten preprint "2+1 Dimensional Gravity as an Exactly Soluable System", Sept. 1988, IASSNS-HEP-88/32]. Thus the "vacuum space-time" is always flat ($R_{abcd}=0$), the (classical) interest coming from global topology and "deficit angles" as regions of "non-vacuum" (singular world-lines?) are circumnavigated.



Twistor theory offers a very different viewpoint as to the possible physical relevance of a TQFT. We are familiar with the idea that there need be no local information stored in a deformed twistor space, or in a twistor space with a holomorphic bundle over it, or a 1-function (element of an H^1) defined on it. Thus, the physically appropriate home for TQFT might well be twistor space rather than space-time.

There is, of course, one fundamental difference between the envisaged setting for TQFT and twistor space: TQFTs refer to <u>real</u> manifolds whereas twistor spaces are <u>complex</u>. However, this need not be a drawback. Ideas can be carried over from the real to the (holomorphic-) complex whenever there is an <u>analytic integral formula</u> expressing what is needed. This can lead to some surprising applications.

Consider the ordinary <u>linking number</u> L for two closed curves in Euclidean 3-space. There is a formula (cf. Flanders, "Differential Forms", p. 80)

$$\mathbf{L} = \frac{1}{4\pi}\oint\!\!\oint \frac{(\vec{x}-\vec{y})\cdot d\vec{x}\wedge d\vec{y}}{|\vec{x}-\vec{y}|^3}$$

for this (expressing the work done on a magnetic monopole carried around one curve, where there is a fixed current in the other — or, no more physically, the work done on an electric charge carried around a "magnetic current"!). We want to interpret this formula as a complex contour integral over (contours in) two holomorphic curves in $\mathbb{CP}^3$, i.e. in $\mathbb{PT}$. (N.B. curves in $\mathbb{RP}^3$ can have quarter-integral linking numbers; but the following might encounter unwanted branching in the $(\ldots)^{3/2}$ in those cases.) We can rewrite L <u>projectively</u> (using the sign $\doteq$ to denote equality up to a multiplicative factor that I haven't yet bothered to work out) as

$$\mathbf{L} \doteq \oint \frac{I\,I\,dx\wedge dy}{(I\,I)^{3/2}} \doteq \oint \frac{I\,I\,dx\wedge dx'\wedge dy\wedge dy'}{(I\,I)^{3/2}}$$

[if we prefer a non-projective version]

where I and I are twistors specifying points on the two curves X and Y (not necessarily compact), respectively, each varying along its appropriate (1-real-dim. closed) contour, and where $\mathcal{Y}$ specifies an arbitrarily chosen "metric" on $\mathbb{PT}$, either complex Euclidean, in which case $(I\,I\,I\,I\,) = 24\det(\mathcal{Y}))$ vanishes, or complexified sphere metric $(\det(\mathcal{Y})\neq 0)$. This expression for L is independent of small finite deformations of

(a) the contours (obviously)
(b) the curves X and Y (not so ")
(c) the "metric" $\mathcal{U}$ (not so ").

These variations can take place within some neighbourhood $\mathcal{N}$ (a complex manifold) of the contours in $\mathbb{PT}$. We must make sure that the "distance" between $\star$ and $\dagger$ never becomes zero (i.e. $\mathcal{U}_{\star\dagger} \neq 0$) within $\mathcal{N}$ (or, rather, on the contours), in our variation.

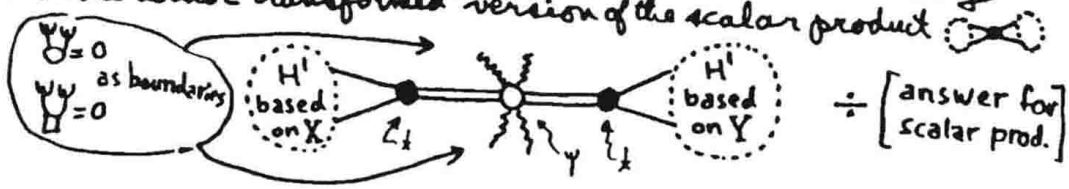
Note the curious fact that this "linking" between X and Y is not simply a topological property since real 2-surfaces in real 6-space don't link. It only works because we are allowing only holomorphic variations of X, Y, $\mathcal{U}$.

The following expression, for which I am grateful for important assistance from A.P.H. → $\mathcal{U}_{\star\dagger}/(\mathcal{U}_{\star\dagger})^{3/2} = \oint \frac{\mathfrak{D}w \log[\Psi\Psi/\Psi\Psi]}{(\star\,\dagger)^2}$ (Ψ "arbitrary") (introduce new Ψ) can be used to express L as:

$$L = \oint_{(\text{proj.})} \frac{\log[\Psi\Psi/\Psi\Psi]\ \Psi d\psi_A d\psi_{\dot A} d\psi \ \star_{\dagger}\, {}_A d\star_{\dot A} d\dagger}{(\star\,\dagger)^2} = \oint_{(\text{non-proj})} \frac{\log[\Psi\Psi/\Psi\Psi]\ d^4w}{(\star\,\dagger)^2}$$

which is handy for examining its invariance properties.

A.P.H. points out that this expression is essentially a "double twistor transformed" version of the scalar product ⟨ ⟩ which is



Here the I-functions are standard cohomology elements specifying the two curves X, Y (cf. R.P. in TN 4 ; M.G.E. & L.P.H. in TN 8 ; both reprinted in Adv. in Tw. Th., for the case of a twisted cubic). M.F.A. recently suggested to me a possible role, in this context, for such scalar product expressions (related to Green's functions), and I am grateful to him for valuable conversations.

Thanks are due also to A.P.H. and other members of the twistor group.

QUESTION: can causality relations between points in M be understood in terms of holomorphic linking of the corresponding $\mathbb{CP}'$s in $\mathbb{PT}$? How does this relate to Rob Low's thesis work?

~ Roger Penrose

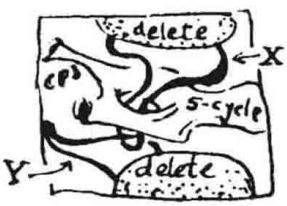
183
HOLOMORPHIC LINKING: POSTSCRIPT

Twistor Newsletter **27** (1988), 4

Holomorphic Linking: Postscript

If we formally integrate out the $\equiv$ in the double twistor transform $(X)\equiv(Y)$ to get $(X)\times(Y)$, and represent each 1-function (X), (Y) locally by a pair of equations $f=0=g$, $h=0=j$, respectively, then we get $\mathbf{L} = \sum\oint \frac{f'(\lambda)\,g'(\lambda)\,h'(\lambda)\,j'(\lambda)\,d^4Z}{f(\lambda)\,g(\lambda)\,h(\lambda)\,j(\lambda)}$ ($x^\alpha = \lambda\pi^\alpha + \xi^\alpha$)

ie. $\mathbf{L} = \sum\oint \frac{df\wedge dg\wedge dh\wedge dj}{fghj} = \sum(\pm(2\pi i)^4)$, where the "$\Sigma$" tots up all the needed pieces for the two 1-functions. In the case where X and Y are algebraic curves, we can use a dot product of the two H^1's to get an H^3 elem. & evaluate this 3-function over $\mathbb{PT}$ to get a canonical value for $\mathbf{L}$ (i.e. indep. of a selection of contours). If we normalize $\mathbf{L}$ so that $\mathbf{L}=1$ for two lines in $\mathbb{CP}^3$, then we get $\boxed{\mathbf{L} = xy}$ where x and y are the orders of X and Y, respectively (number of points they meet a generic plane). This is one quarter of the ("standard") linking number which would be 1 for two linking circles. (This result can be obtained by specializing X to x lines and Y to y lines.) A possible suggestion for the more general holomorphic linkings (e.g. distinguishing $\boxed{OO}$ from $\boxed{\text{⊂⊃}}$!) might be to find a real 5-cycle separating X from Y (in $\mathbb{CP}^3$) and such that $\exists$ 1-functions defining X and Y whose domains intersect in a region containing the 5-cycle. Each 1-function is a 1-form projectively and a 2-form non-projectively. Use the non-projective form, and cup product to obtain a (2+2)-form (1+1)-function, i.e. element of $H^2(d\Omega^3)$ (since closed). We can evaluate this on the non-projective 6-cycle which is the circle bundle over the 5-cycle in the S^7 of "unit circles" over $\mathbb{CP}^3$ (in $\mathbb{C}^4$). I don't know if this gives significantly more than we had before.



REMARK I find the relation between holomorphic linking and twistor diagrams intriguing. Linking (at least, real linking) is a combinatorial thing. Could there be a deep combinatorial aspect to twistor diagram theory? Could this relate to the original motives lying behind spin-network theory?

– Roger Penrose

184
SINGULARITIES AND BIG-BANG COSMOLOGY

Q. J. R. Astr. Soc. **29** (1988), 61–3

Singularities and Big-Bang Cosmology

R.Penrose
Mathematical Institute, Oxford

(Received 1987 September 16)

One of the more disquieting features of the cosmological solutions of the equations of Einstein's general relativity, as put forward by Friedmann, Lemaître, Robertson, Walker and others, was the existence of a singular origin to the Universe: the big bang. Various people speculated – and Einstein among them – that the presence of such an initial singularity might be a feature merely of the high symmetry that was assumed for these models. Thus, so it had seemed, what appeared to be a singular origin for the world could perhaps be merely a highly compressed and irregular *non*-singular state. We could imagine this state being the result of a previous collapsing phase of the Universe, where the presence of asymmetries – magnified in their effect to an enormous degree at the moment of maximum compression – might allow an actual space–time singularity to be avoided. In apparent agreement with such a conclusion, the work of Lifshitz & Khalatnikov (1963) had seemed to imply that singularities would simply not arise in generic situations. Accordingly, the collapsing universe of the previous phase would not become singular, but would be seen to 'bounce out again' and provide a not-quite-singular big bang, consistent with what we seem to perceive for our own Universe.

Also, on this view, the singularities in the black holes which arise in models of exactly symmetrical gravitational collapse (in the Schwarzschild solution; cf. Oppenheimer & Synder 1939) would similarly be eliminated if generic asymmetries were incorporated. However, this turned out not to be the case. In the early 1960s there had been a renewed interest in such questions in the relativity community – most particularly on the part of John Wheeler – an interest which had resulted from the first quasar observations. This had stimulated me, in late 1964, to re-examine a question that had intrigued me since my first serious involvement in general relativity, namely: is the Schwarzschild singularity (at $r = 0$) actually stable? By employing some differential–topological techniques that I had been developing for a different purpose (cf. appendix to Penrose 1965a) I was able to obtain a general result which showed, in effect, that generic gravitational collapse, according to standard Einstein theory, will inevitably result in space–time singularities (Penrose 1965b). Thus the presence of asymmetries will not save us from singularities, and the Schwarzschild singularity indeed has a certain stability about it – in the sense that perturbations do not remove space–time

* Based on a talk given at the RAS Specialist Discussion on *Astronomy in Britain since the Second World War*, 1987 May 8, organized by Mr Colin Ronan.

singularities altogether (though in a more *local* sense the Schwarzschild singularity actually turns out to be unstable).

By turning the argument 'upside down' Hawking (1965) then showed that it could be applied also to the Big Bang. He then refined the techniques to enable various improved results to be obtained which removed certain possible loopholes (Hawking 1966a, b, 1967); and, finally, we combined forces to produce a comprehensive theorem on this topic (Hawking & Penrose 1970). In the meanwhile, Belinskii had joined Lifshitz and Khalatnikov; and, finding a technical error in their previous work, they came to a conclusion which was in accord with our own (Belinskii, Khalatnikov & Lifshitz 1970).

An important question remained, however; are the singularities that result from gravitational collapse necessarily hidden from the view of an outside observer by an *event horizon*, or is it possible that sometimes *naked singularities* might arise, which might have direct astrophysical implications of an alarming kind? This is the question of *cosmic censorship* (Penrose 1969). Is there some profound implication of the equations of classical general relativity that forbids naked singularities arising from generic collapse? The general question remains unanswered – and is perhaps the most important unsolved problem of classical general relativity theory. The balance of informed opinion on this question seems to have swung fairly firmly to the view that such a cosmic censorship principle probably does hold, but no proof is at hand.

If a rather strong version of cosmic censorship holds (implying that an observer inside a black hole will not even see the singularity before being absorbed by it) then space–time singularities can be clearly divided into two kinds: *past* singularities and *future* singularities (Penrose 1979). There is impressive evidence – from the second law of thermodynamics as well as from the powerful observational facts of cosmology – that past and future singularities are completely different from one another in detailed geometric structure. The future singularities would have an enormously high entropy, while for the big bang some huge constraint must apply, reducing the entropy to a comparatively tiny value. I have suggested elsewhere (cf. Penrose 1979) that this difference in singularity structure can be encapsulated in the hypothesis that the *Weyl conformal curvature* must *vanish* in the immediate neighbourhood of past singularities, but that no such constraint applies to future singularities. Assuming, as seems reasonable, that the physical rules governing the structure of the classical space–time singularities come from *quantum gravity* – whatever the correct union of quantum theory with general relativity ultimately turns out to be – we seem to have the implication that quantum gravity must be a *time-asymmetric* theory! (This implication is quite unaffected by current considerations of 'inflationary universes'. Such inflationary considerations in no way remove the necessity for an absolutely stupendous constraint to be placed on the initial singular state.) In my opinion (cf. Penrose 1987), there are strong indications that considerations of quantum gravity will eventually yield important changes in the actual structure of quantum theory, not just in general relativity!

REFERENCES

Belinskii, V.A., Khalatnikov, I.M. & Lifshitz, E.M., 1970. *Adv. Phys.* **19**, 523.
Hawking, S.W., 1965. *Phys. Rev. Lett.* **15**, 689.
Hawking, S.W., 1966a. *Proc. R. Soc.* A **294**, 511.
Hawking, S.W., 1966b. *Proc. R. Soc.* A **295**, 490.
Hawking, S.W., 1967. *Proc. R. Soc.* A **300**, 187.
Hawking, S.W. & Penrose, R., 1970. The singularities of gravitational collapse and cosmology, *Proc. R. Soc.* A **314**, 529.
Lifshitz, E.M. & Khalatnikov, I.M., 1963. *Adv. Phys.* **12**, 185.
Oppenheimer, J.R. & Synder, H., 1939. *Phys. Rev.* **56**, 455.
Penrose, R., 1965a. Zero rest-mass fields including gravitation: asymptotic behaviour, *Proc. R. Soc.* A **284**, 159.
Penrose, R., 1965b. Gravitational collapse and space-time singularities, *Phys. Rev. Lett.* **14**, 57.
Penrose, R., 1969. Gravitational collapse: the role of general relativity, *Rivista del Nuovo Cimento*: *Numero speciale* **1**, 252.
Penrose, R., 1979. Singularities and time-asymmetry, in *General Relativity*: *an Einstein Centenary*, eds Hawking, S.W., & Israel, W., Cambridge University Press.
Penrose, R., 1987. Newton, quantum theory and reality, in *300 Years of Gravity*, eds Hawking, S.W., & Israel, W., Cambridge University Press.

185
ON THE PHYSICS AND MATHEMATICS OF THOUGHT

The universal Turing machine: A half-century survey, 491–522, Oxford University Press, New York, 1988. Also published in: *Logic, Philosophy of Science and Epistomology.* Proc. 11th Int. Wittgenstein Symp., 122–8, Hölder-Pichler-Tempskey, Wien; Kammerer and Unverzagt (1998); and Springer 1994/95

Much of this article was lifted from the parts of my book *The Emperor's New Mind* (Oxford University Press, 1989) that I had written at the time of my contributing to this volume, honouring Alan Turing. Now Gödel's results do begin to play their role (compare Chapter 170).

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey*, 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

On the Physics and Mathematics of Thought

Roger Penrose

1. Introduction

The questions that I shall touch upon here are closely related to those that Alan Turing concerned himself with in the seminal article that this volume commemorates (Turing *1936-7*). What are the essential physical actions that take place when a person thinks? Is a human brain, indeed, effectively a universal Turing machine? Is the precise material that constitutes a thinking device (such as a brain) of any real consequence? Does it much matter what physical laws are called upon in the action of such a device?

Regrettably I have not time, as I write this, to develop fully the (rather incomplete) ideas that I would wish to express; nor is there really the necessary space for an adequate discussion. As it happens, I am in the process of writing a semi-popular but lengthy account of those ideas, which will be published in due course by Oxford University Press (under the title: *The Emperor's New Mind*). Instead of attempting to summarize the intended contents of that work here, I shall present an abbreviated account based largely on a few excerpts from portions that are already written, albeit in their present preliminary state. (I am grateful to Oxford University Press for their permission to proceed in this way.)

2. Is Mental Activity Merely the Enacting of an Algorithm?

A certain view seems to have gained prominence in recent years, namely that a human brain is, indeed, effectively some kind of computer, and that it differs from a universal Turing machine only in that its storage capacity is finite, that its activity is curtailed in time, that it may involve some random elements, and that it sometimes makes mistakes. This, indeed, seems to

have been Turing's own view, more or less (Turing *1950*), and his general concept of what could constitute a thinking device—in particular, a brain—led him originally to his powerful Turing machine concept. On this view, it does not *in principle* matter how the device is physically constructed. Cogs, wires and pulleys, water pipes, or electromagnetic waves would each do as well as electronics, or the electro-chemical activity of neurones. Though the details of the physical action might be different in each case, this would be of no overall consequence—apart from the (perhaps very great) differences in capacity, speed, and reliability.

Must we accede to this view? Indeed, is it necessarily the case that any physical object whatever—where I include a functioning human brain as a physical object—must act by rules that are in accordance with the running of some Turing machine program? Or may it be that some suitably delicately organized (perhaps biological) construction can perform in some fundamentally different, and perhaps even superior, way? My own opinion, for what it is worth, is that this latter possibility is very likely to be true of, in particular, the functioning of (parts of) our brains, and that there is some noncomputable element present in physical laws which may well be being harnessed in the physical basis of our thoughts.

This is, I believe, a distinctly unconventional view. Turing had, after all, presented a very persuasive case that any mechanism, in the ordinary sense of the word, would have to follow the general procedures that he laid down for a Turing machine. I am not disputing his line of argument, which I, also, find persuasive. What I am attempting to dispute is the idea that a physical object need actually behave as a "mechanism" in this ordinary sense. As part of my argument I shall need to explain what I mean by a physical object not behaving as a mechanism. For the moment, let us simply say that such a physical "nonmechanism" might have a behavior which is completely governed by precise physical laws, yet governed in a way that cannot be derived from any algorithm.

My reasons for adopting such a viewpoint come from several different directions. One of these has to do with our conscious perceptions. To investigate this, let us first consider a profoundly opposed viewpoint—that of "strong AI"—which asserts that our thinking consists merely of the enacting of some algorithm. According to this view, consciousness must be some quality that can be actually possessed by algorithms. Somehow, this quality would have to become manifest whenever that algorithm is enacted. However, John Searle (*1980*) has argued strongly that the mere enacting of an algorithm cannot, in itself, evoke the mental qualities that that algorithm is intended to provide. He presents his example of a "Chinese room", in which a human subject manipulates counters with Chinese symbols marked on them—these manipulations following the action of some

On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

known algorithm—but without experiencing the mental quality, in this case "understanding", that the algorithm is supposed to provide. Searle's arguments are, in my opinion, interesting and fairly persuasive, but by no means definitive.

There is another fundamental problem for the strong AI position: what does it actually *mean* for an algorithm to be "enacted"? Does "enacting" necessarily involve the movement of some material bodies (e.g., of electrons along wires, or of neurological chemicals) in some spatio-temporal configuration which represents the *running* of the algorithm? Or can "mental qualities" be achieved merely by the algorithm's specification having been written out on some page (or in the magnetic fields of some computer disc)? At what stage would "conscious awareness" be conjured into being, according to this view? If that conscious awareness requires the physical movement of bodies, then it is *not* merely a quality possessed by algorithms: conscious thinking would indeed depend upon the *physical action* of the "thinking device" that is involved. If the mere specification of an algorithm can conjure up "awareness", then why does it need physical embodiment at all? Perhaps the algorithm's mere Platonic existence as a mathematical object is, in itself, sufficient to give it conscious feelings!

This dilemma lies behind the scenes of an argument put forward by Douglas Hofstadter (*1981*)—himself a major proponent of the strong-AI view—in a dialogue entitled "A Conversation with Einstein's Brain". Hofstadter envisages a book, of absurdly monstrous proportions, which is supposed to contain a complete description of the brain of Albert Einstein. Any question that one might care to put to Einstein can be answered just as the living Einstein would have, simply by leafing through the book and carefully following all the detailed instructions it provides. Of course "simply" is an utter misnomer, as Hofstadter is careful to point out. But his claim is that *in principle* the book is completely equivalent, in the operational sense of a "Turing test", to a ridiculously slowed down version of Einstein himself. (Recall Turings's idea of a test whereby an interrogator has to decide, merely from the answers to a succession of questions, which of two subjects is a human being and which is a computer. If she consistently cannot tell, then the computer passes, and is deemed to possess the appropriate "mental qualities". We imagine a corresponding criterion for deciding, in effect, whether something—here a book—"is" or "is not" Einstein!) Thus, according to the contentions of strong AI, the book would think, feel, understand, be aware—just as though it were Einstein himself, but no doubt living at a monstrously slowed-down rate (so that to the book-Einstein the world outside would seem to flash by at a ridiculously speeded-up rate). Indeed, since the book is supposed to be merely a particular embodiment of the algorithm which constitutes Einstein's "self", it would actually *be* Einstein.

But now a difficulty presents itself. The book might never be opened, or might be continually pored over by innumerable students and searchers after truth. How would the book "know" the difference? Perhaps the book would not need to be opened, its information being retrieved by means of X-ray tomography, or some other technological wizardry. Would Einstein's awareness be enacted only when the book is being so examined? Would he be aware twice over if two people chose to ask the book the same question at two completely different times? Or would that entail two separate and temporally distinct instances of the *same* internal state of Einstein's awareness? Perhaps his awareness would be manifested only if the book is *changed*? After all, normally when we are aware of something we receive information from the outside world which affects our memories—and the states of our minds are indeed slightly changed. If so, does this mean that it is (suitable) *changes* in algorithms (and here I am including the memory store as part of the algorithm) which are to be associated with mental events rather than (or perhaps in addition to) the *activation* of algorithms? Or would the book-Einstein remain completely self-aware even if it were never examined or disturbed by anyone or anything? Hofstadter touches on (some of) these questions, but he does not really attempt to answer them, nor even properly to come to terms with most of them.

What does it mean to activate an algorithm, or to embody it in physical form? Is changing an algorithm different in any sense from merely discarding one algorithm and replacing it with another? What on earth does any of this have to do with our feelings of conscious awareness? The reader (unless himself or herself a supporter of strong AI) may be wondering why I have devoted this much space to such a patently absurd idea. In fact, I do *not* regard the idea as intrinsically an absurd one—mainly just wrong!

3. *Is All Physical Action Algorithmic?*

There is, indeed, a certain force in the reasoning behind strong AI which must be reckoned with. If, as appears to be the common belief, it is actually true that the action of *any* physical object is equivalent to the running of some Turing machine, then in particular it must be the case that the activity of a human brain can, in principle, be exactly modeled by a general-purpose computer (i.e., in effect, by a universal Turing machine). Any question that one might put to the human could, in principle, be answered in an identical manner by the computer. Thus, if one adopts Turings's highly *operational* viewpoint—as described according to his "Turing test"—one must accept that the computer, whenever it is (in an appropriate sense) enacting the

appropriate algorithm, must have the same mental experiences as the human would.

It may be argued that some such operational procedure is, after all, the way that one judges the presence of mental qualities in human beings other than oneself. The strong-AI viewpoint merely extends this type of judgement from biological structures such as ourselves to other (e.g., electronic) kinds of device. In my own opinion, there is a certain rationale to this. I do not myself see why biological systems should be singled out as the only objects capable of evoking mental qualities. However, I do feel that there is a *different* unjustified assumption being invoked here, namely that all objects in the physical world must necessarily behave algorithmically.

Even Searle, in some of his discussions, seems to be implicitly accepting that electronic computers of the present-day type, but with considerably enhanced speed of action and size of rapid-access store (and possibly parallel action) may well be able to pass the Turing test proper, in the not-too-distant future. He is prepared to accept the contention of strong AI (and of most other "scientific" viewpoints) that "we are the instantiations of any number of computer programs" (Searle *1980,* in Hofstadter and Dennett *1981,* p. 368); moreover, he succumbs to: "Of course the brain is a digital computer. Since everything is a digital computer, brains are too" (ibid., p. 372). Searle maintains that the distinction between the functioning of human brains (which can have minds) and of electronic computers (which, he has argued, cannot) both of which might be executing the same algorithm, lies solely in the material construction of each. He claims, but for reasons he is not able to explain, that the biological objects (brains) can have "intentionality" and "semantics", which he regards as defining characteristics of mental activity, whereas the electronic ones cannot. Just in itself this does not seem to me to point the way towards any helpful scientific theory of mind. What is so special about biological systems, apart perhaps from the "historical" way in which they have evolved (and the fact that *we* happen to be such), which sets them apart as the objects allowed to achieve intentionality or semantics? The claim looks to me suspiciously like a dogmatic assertion, perhaps no less dogmatic, even, than those assertions of strong AI which maintain that the mere enacting of an algorithm can conjure up a state of conscious awareness!

In my opinion Searle, and a great many other people, have been led astray by the computer people. And they, in turn, have led astray by the physicists. (It is not the physicists' fault. Even *they* don't know everything!) The belief seems to be widespread that, indeed, "everything is a digital computer". It is my intention, eventually, to try to show why, and perhaps how, this need *not* be the case. I shall not be able to develop the full arguments here, however. These must await the completion of my more extended account,

but I shall present some indication of them later in this article. For the moment, suppose we accept the commonly believed view that all physical action is indeed algorithmic. We shall see why the strong-AI viewpoint is in some ways a natural development from this.

4. *Hardware and Software*

In the jargon of computer science, the term *hardware* is used to denote the actual machinery involved in a computer (printed circuits, transistors, wires, etc.), including the complete specification for the way in which everything is connected up. Correspondingly, the term *software* refers to the various programs which can be run on the machine. It was one of Turing's remarkable discoveries that, in effect, any machine whatever, for which the hardware has achieved a certain definite degree of complication and flexibility—that of a *universal* Turing machine—is *equivalent* to any other such machine. This equivalence is to be taken in the sense that for any two such machines A and B there would be a specific piece of software which if given first to machine A would make it act precisely as though it were machine B and, likewise, there would be another piece of software which would make machine B act precisely like machine A. I am using the word "precisely" here to refer to the actual output of the machines for any given input (fed in after the converting software is fed in) and *not* to the *time* that each machine might take to produce that output. I am also allowing that if either machine at any stage runs out of storage space for its calculations then it can call upon some (in principle unlimited) external supply of blank "rough paper"—which could take the form of magnetic tape, discs, drums, or whatever. In fact, the difference in time taken over some task, between the machines A and B, might well be a very serious consideration. It might be the case, for example, that A is much more than a thousand times faster at performing a particular task than is B. It might also be the case that, for the very same machines, there is some other task for which B is over a thousand times faster than A. Moreover, these timings could depend very greatly on the particular choices of converting software that are used. This is very much an "in principle" discussion, where one is not really concerned with such practical matters as achieving one's calculations in a reasonable time.

Since, in effect, all modern general purpose computers are universal Turing machines, they are equivalent to one another in the above sense: the differences between them can be entirely subsumed in the software—provided that we are not concerned about differences in the resulting speed of operation and possible limitations on storage size. Indeed, modern technology

has enabled computers to perform so swiftly and with such vast storage capacities that, for a good many "everyday" purposes, neither of these practical considerations actually represents any serious limitation to what is normally needed, so this effective theoretical equivalence between computers can be seen also at the practical level. Technology has, it seems, transformed entirely academic discussions concerning idealized computing devices into matters which directly affect all of our lives!

As far as I can make out, one of the most important factors underlying the strong-AI philosophy is this equivalence between physical computing devices. The hardware is seen as being relatively unimportant (perhaps even totally unimportant) and the software, i.e., the program, or the algorithm, is taken to be the one vital ingredient. However, it seems to me that there are also other important underlying factors lending some support to this view, these coming more from the direction of physics. I shall next try to give some indication of all this, as I see it.

What is it that gives a particular person his (or her) individual identity? Is it, to some extent, the very atoms that compose his body? Is his identity dependent upon the particular choice of electrons, protons, and other particles that compose those atoms? There are at least two reasons which tell us that this cannot be so. In the first place, there is a continual turnover in the material of any living person's body. This applies in particular to the cells in a person's brain—despite the fact that no new actual brain cells are produced after birth. The vast majority of atoms in each living cell (including each brain cell) are continually being replaced.

The second reason comes from quantum physics—and by a strange irony is, strictly speaking, in contradiction with the first! According to quantum mechanics, any two electrons must necessarily be completely identical with one another, and the same holds for any two protons and for any two particles whatever, of any one particular kind. This is not merely to say that there is no way of telling the particles apart. The statement is considerably stronger than that. If an electron in a person's brain were to be exchanged with an electron in a brick, then the state of the system would be *exactly the same state* as it was before, not merely indistinguishable from it! (The knowledgeable reader who is worrying about an unimportant *sign* difference should imagine that one of the electrons is rotated completely through 360° as we make the interchange!) The same holds for protons and for any other kind of particle, and for whole atoms, molecules, etc. If the entire material content of a person were to be exchanged with corresponding particles in the bricks of his house then, in a strong sense, nothing would have happened whatsoever. What distinguishes the person from his house is the *pattern* of how his constituents are arranged, not the individuality of the constituents themselves.

There is perhaps an analogue of this at an everyday level which is independent of quantum mechanics, but made particularly manifest to me as I write this, by the electronic technology which enables me to type at a word processor. If I desire to change a word, say to transform "make" into "made" I may do this by simply replacing the "k" with "d", or I may choose instead to type out the whole word again. If I do the latter, is the "m" the same "m" which was there before, or have I replaced it with another identical one? What about the "e"? Even if I do simply replace the "k" with "d", rather than retype the word, there is a moment just between the disappearance of "k" and reappearance of "d" when the gap closes and there is (or, at least, sometimes is) a wave of realignment down the page as the placement of every succeeding letter (including the "e") is recalculated, and then re-recalculated as the "d" is inserted. (Oh, the cheapness of mindless calculation in this modern age!) In any case, *all* the letters that I see before me on the screen are mere gaps in the track of an electron beam as the whole screen is scanned some fifty times each second. If I take any letter whatever and replace it with an identical one, is the situation the *same* after the replacement, or merely indistinguishable from it? To try to adopt the second viewpoint (i.e., "merely indistinguishable") as being distinct from the first (i.e., "the same") seems footling. At least, it seems reasonable to call the situation the same when the letters are the same. And so it is with the quantum mechanics of identical particles. To replace one particle by an identical one is actually to have done nothing to the state at all.

The earlier remarks above concerning the continual turnover of atoms in a person's body were made in the context of classical rather than quantum physics. The remarks were worded as though it might be meaningful to maintain the individuality of each atom. In fact classical physics is adequate for this, and we do not go badly wrong, at this level of description, by regarding atoms as individual objects. Provided that the atoms are reasonably well separated from their identical counterparts as they move about, one *can* consistently refer to them as maintaining their individual identities since each atom can be, in effect, tracked continuously, so that one could envisage keeping a tab on each separately. From the point of view of quantum mechanics it would be a convenience of speech only to refer to the individuality of the atoms, but it is a consistent enough description at the level just considered.

Let us accept that a person's individuality has nothing to do with any individuality that one might try to assign to his material constituents. Instead, it must have to do with the *configuration,* in some sense, of those constituents—let us say the configuration in space or in space-time. But the supporters of strong AI go further than this. If the information content of such a configuration can be translated into another form from which the

original can again be recovered then, they claim, the person's individuality must remain intact. It is like the sequences of letters I have just typed and now see displayed on the screen of my word processor. If I move them off the screen, they remain coded in the form of certain tiny displacements of electric charge, in some configuration in no clear way geometrically resembling the letters I have just typed. Yet, at any time I can move them back on the screen, and there they are, just as though no transformation had taken place. If I choose to save the part I have just written, then I can transfer the information of the sequences of letters into configurations of magnetization on a disc which I can later remove, and by switching off the machine I neutralize all the (relevant) tiny charge displacements in it. Tomorrow, I can reinsert the disc, reinstate the little charge displacements, and display the letter sequences again on the screen, just as though nothing had happened. To the strong-AI supporters, it is "clear" that a person's individuality can be treated in just the same way. Like the sequences of letters on my display screen, they would claim, nothing is lost of a person's individuality—indeed nothing would really have happened to it at all—if his (or her) physical form were to be translated into something quite different, say into fields of magnetization in a block of iron. They are even driven to claim that the person's conscious awareness would persist while the person's "information" is in this other form. In this view, a "person" is to be taken, in effect, as a piece of software, and his particular manifestation as a material human being is to be taken as the operation of this software by the hardware of his body.

It seems that the reason for these claims is that, whatever material form the hardware takes—for example some electronic device—one could always "ask" the software questions (in the manner of a Turing test), and assuming that the hardware performs satisfactorily in computing the replies to these questions, these replies would be identical to those that the person would make whilst in his normal state. ("How are you feeling this morning?" "Oh, fairly well, thank you, though I have a slightly bothersome headache." "You don't feel, then, that there's ... er ... anything odd about your personal identity ... or something?" "No; why do you say that? It seems to be a rather strange question to be asking." "Then you feel yourself to be the same person that you were yesterday?" "Of course I do!")

5. *Mathematical Platonism*

Let us try to go along, for a while, with the strong-AI view that the algorithm (or software) is all that counts. Do algorithms have some kind of existence on their own as mathematical objects, independent of any particu-

lar physical form? This question returns us to the matter of the Platonic existence of mathematical objects—in particular, of algorithms—and the issue of whether such "existence" might, according to the strong-AI philosophy, be in itself sufficient to give an algorithm conscious feelings! One is led, here, into questions of the philosophy of mathematics. Do mathematical concepts have any (Platonic) existence beyond their mere presence in the minds of mathematicians? If mathematical concepts could exist only in minds, then it would not be of much use attempting to explain minds in terms of algorithms—using a concept of an algorithm which can itself only exist in a mind! One might take the view that mathematical concepts could also exist in terms of some other physical embodiment, such as on the printed page or in a computer memory. But there is the difficulty here that such physical embodiments depend upon the language in which they are written. This language must be "known" in order that the mathematical concept being described be actually well defined—but "known" by whom? It seems that one is led back to the necessary presence of a mind in order to do the "knowing", and one is no better off than before.

To avoid such circularity, it seems that strong-AI supporters must take a more Platonic view of mathematics than that. Let us now address this question, and try to examine how plausible the Platonic view is. How "real", indeed, *are* the objects of the mathematician's world? From one point of view it seems that there can be nothing real about them at all. Mathematical objects are just concepts; they are the mental idealizations that mathematicians make, often stimulated by the appearance and seeming order of aspects of the world about us, but mental idealizations nevertheless. Can they be other than mere arbitrary constructions of the human mind? On the other hand, there often does appear to be some profound reality about these mathematical concepts, going quite beyond the mental deliberations of any particular mathematician. It is as though human thought is, instead, being guided towards some eternal external truth—a truth which has a reality of its own, and which is revealed only partially to any one of us.

The familiar Mandelbrot set provides a striking example. Its wonderfully elaborate structure was not the invention of any one person—nor was it the design of a team of mathematicians. Benoit Mandelbrot himself, who first studied the set, had no real prior conception of the fantastic elaboration inherent in it, although he certainly knew that he was on the track of something very interesting. Indeed, when his first computer pictures began to emerge, he was under the impression that the fuzzy structures that he was seeing were the result of a computer malfunction (Mandelbrot *1986*)! Only later did he become convinced that they were really there in the set itself. Moreover, the complete details of the complication of the structure of Mandelbrot's set cannot really be fully comprehended by any one of us, nor

can it be fully revealed by any computer. It would seem that this structure is not just part of our minds, but that it has a reality of its own. Whichever mathematician or computer buff chooses to examine the set, approximations to the *same* fundamental mathematical structure will be found. It makes no real difference which computer is used for performing calculations (provided that the computer is in accurate working order), apart from the fact that differences in computer speed and storage, and graphic display capabilities, may lead to differences in the amount of fine detail that will be revealed and in the speed with which that detail is produced. The computer is being used in essentially the same way that the experimental physicist uses a piece of experimental apparatus to explore the structure of the physical world. The Mandelbrot set is not an invention of the human mind. It was a discovery. Like Mount Everest, the Mandelbrot set is just *there*!

Likewise, the very system of complex numbers, in terms of which that set is described, has a profound and timeless reality which goes quite beyond the mental constructions of any particular mathematician. The beginnings of an appreciation of complex numbers came about in the sixteenth century with the work of Gerolamo Cardano. In this he put forward the first complete expression for the solution (in terms of surds) of a general cubic equation. He noticed, however, that in a certain class of cases—the ones referred to as "irreducible", where the equation has three real solutions—he was forced to take, at a certain stage in his expression, the *square root of a negative number.* Although this was puzzling to him, he realized that if he allowed himself to *take* such square roots, and *only* if so, then he could express the full *real* answer!

While at first it may seem that the introduction of such roots of negative numbers was just a device—a mathematical invention designed to achieve a specific purpose—it later becomes clear that these objects are achieving far more than that for which they were originally designed. Although the original purpose of introducing complex numbers was to enable square roots to be taken with impunity, by introducing such numbers we find that we get, completely free, the potentiality for taking any other kind of root or of solving any algebraic equation whatever. Later we find many other magical properties that these complex numbers possess—properties that we had no inkling about at first. These properties are just *there.* They were not put there by Cardano—nor by Bombelli, nor Wallis, nor Cotes, nor Euler, nor Wessel, nor Gauss, despite the undoubted farsightedness of these and other great mathematicians—but such magic was inherent in the very structure that they gradually uncovered. When introducing his complex numbers, Cardano could have had no inkling of the many other magical properties which were to follow—properties which go under various names, such as: the Cauchy integral formula, the Riemann mapping theorem, the Lewy ex-

tension property. These, and many other equally remarkable facts, are properties of the very numbers, with no additional modifications whatever, that Cardano had first encountered in about 1539.

All this suggests that the finding of mathematical results is a matter of *discovery* and not invention. Yet the matter is perhaps not quite so straightforward as this. As I have said, there are things in mathematics for which the term "discovery" is indeed much more appropriate than "invention", such as the examples just cited. These are the cases where much more comes out of the structure than is put into it in the first place. One may take the view that in such cases the mathematicians have stumbled upon "works of God". However, there are other cases where the mathematical structure does not have such a compelling uniqueness, such as when in the midst of a proof of some result, the mathematician may find the need to introduce some complicated and far from unique construction in order to achieve one very specific end. In such cases no more is likely to come out of the construction than was put into it in the first place, and the word "invention" seems more appropriate than "discovery". These are indeed just "works of man". On this view, the true mathematical discoveries would, in a general way, be regarded as greater achievements or aspirations than would the "mere" inventions.

Such categorizations are not entirely dissimilar from those that one might use in the arts or in engineering. Great works of art are indeed "closer to God" than are lesser ones. It is a feeling not uncommon amongst artists that in their greatest works they are revealing eternal truths which have some kind of prior ethereal existence, whilst their lesser works might be more arbitrary, of the nature of mere mortal constructions. Likewise, an engineering innovation with a beautiful economy, where a great deal is achieved in the scope of the application of some simple unexpected idea, might appropriately be described as a discovery rather than an invention.

Having made these points, however, I cannot help feeling that, with mathematics, the case for believing in some kind of external existence, at least for the more profound mathematical concepts, is a good deal stronger than in those other cases. There is a compelling uniqueness and universality in such mathematical ideas which seems to be of quite a different order from that which one could expect in the arts or engineering.

How does this relate to the issues of strong AI, according to which mental phenomena are supposed to find their existence within the mathematical idea of an algorithm? In my opinion, the actual concept of an algorithm is indeed a profound and "God-given" notion. Does this viewpoint lend some credence to the strong AI point of view, by providing the possibility of an ethereal type of existence for mental phenomena? Just conceivably so; but if mental phenomena can indeed find a home of this general kind, I do not

believe that it can be with the concept of an algorithm. For one thing, strong AI attributes mental significance to *specific* algorithms rather than to be the more "God-given" concept of a *general* algorithm. What would be needed would be something very much more subtle than specific algorithms. The fact that algorithmic things indeed constitute a very narrow and limited part of mathematics will be an important aspect of the later discussions.

6. *Gödel's Theorem*

Are there more clear-cut reasons than these for doubting that conscious thought is merely the enacting of an algorithm? I believe that there are. In particular, it would appear that our (conscious) perception of mathematical truth goes demonstrably beyond anything that can be formulated in algorithmic terms. Gödels theorem seems to show this, since Turing effectively established that the algorithmic decidability of an area of mathematics is equivalent to its axiomatizability, according to Hilbert's program. Gödel's result tells us that our perceptions of mathematical truth indeed cannot be axiomatized. (Compare Lucas *1961.* Counterarguments to Lucas's line of reasoning have often been put forward, e.g., Benacerraf *1967,* Good *1969,* Bowie *1982,* Hofstadter *1979.* In a general way, I am in sympathy with the Lucas type of viewpoint, though my own viewpoint is perhaps a little different. I do not think that the arguments are conclusive either way. Some of Lucas's critics seem to make the assumption that the behavior of any physical object—e.g., Lucas himself—must be governed by *some* algorithm. It is this assumption that I am disputing!)

Let us briefly recall the Hilbert program: each clear-cut body of mathematical thought is to be encapsulated in a system of axioms and rules of procedure. The axioms are to be expressed as strings of symbols belonging to some precisely defined alphabet, the rules of procedure being some clear set of allowed ways of deriving new such strings of symbols from the ones already obtained. But the hopes of David Hilbert and his followers were dashed by Kurt Gödel in 1931. Gödel showed that any such precise ("formal") mathematical system of axioms and rules of procedure *whatever,* provided that it is broad enough to contain descriptions of simple arithmetical propositions (like "Fermat's last theorem") and provided that it is free from contradiction, must contain some statements which are neither provable nor disprovable by the means allowed within the system. It will be important for us to understand the nature of this "undecidability" and see why Gödel's argument cut to the very core of the Hilbert program. We must see how Gödel's argument indeed enables us, by the use of conscious

contemplation and insight, to go beyond the limitations of any particular formalized mathematical system.

According to Hilbert's program, it is in principle possible to check, within such a formalized system, whether any mathematical proof has been correctly given. The hope had been that such a system could be *complete,* and this would have enabled one in principle to decide the truth or falsity of any mathematical statement that could be formulated within the system. Thus for any correctly formed proposition, say, P, one should be able to prove either P or else $\sim P$ (where "$\sim P$" stands for "not P"), depending upon whether P is true or false. In fact, this would even enable us to dispense with worrying about what the propositions *mean* altogether! P would just be a string of symbols, correctly assembled according to certain prescribed ("syntactical") rules (i.e., brackets paired off correctly, etc.). The string of symbols P would be assigned the truth-value *"true"* if P is a theorem (i.e., if P is provable within the system) and it would be assigned the truth-value *"false"* if, on the other hand, $\sim P$ is a theorem. For this to make sense, we require *consistency* in addition to completeness. That is to say, there must be no string of symbols P for which *both* of P and $\sim P$ are theorems. Otherwise P could be *"true"* and *"false"* at the same time.

The point of view that one can dispense with the meanings of mathematical statements, and just regard such statements as being nothing but strings of symbols in some formal mathematical system, is the mathematical standpoint of *formalism.* Some people like that idea. However, it is not one that appeals to me at all. Fortunately, Gödel dealt it a devastating blow! Let us run through his argument to see how he did this.

Part of Gödel's argument was very detailed and complicated, but it is not necessary for us to examine the intricacies of that part. The central idea, on the other hand, was simple, beautiful, and profound. That part we shall be able to appreciate. The complicated part (which also contained much ingenuity) was to show in detail how one may actually code the individual rules of procedure of the formal system, and also the use of its various axioms, into *arithmetical operations.* (It was an aspect of the profound part, though, to realize that this was a fruitful thing to do!) In order to carry out this coding we need to find some convenient way of labeling propositions with natural numbers. One way would be simply to use some kind of "alphabetical" ordering for all the strings of symbols of the formal system for each specific length, where there is an overall ordering according to the length of the string. (Thus, the strings of length one could be alphabetically ordered, followed by the strings of length two, alphabetically ordered, followed by the strings of symbols of length three, etc.) This amounts to using the notation for the strings of symbols in the formal system as a notation for the *natural numbers* expressed in a "base" given by the number

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey*, 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

of symbols of the system. (In fact Gödel originally used a rather more complicated system.) We shall be particularly concerned with (syntactically correctly formulated) *propositional functions* (propositions dependent on numerical variables) which are dependent on a *single* variable. Let the n-th such propositional function (in the chosen ordering of strings of symbols), applied to w, be

$$P_n(w).$$

Then $P_n(w)$ will be some perfectly well-defined particular arithmetical statement concerning the two natural numbers n and w. Precisely *which* arithmetical statement it is will depend on the details of the particular numbering system that has been chosen. That belongs to the complicated part of the argument and will not concern us here. The strings of propositions which constitute a *proof* of some theorem in the system can also be labeled by natural numbers using the chosen ordering scheme. Let

$$\Pi_n$$

denote the n-th proof.

Now consider the following propositional function, which depends on the natural number w (where "$\exists$" means "there exists ..."):

$$\sim \exists x[\Pi_x \text{ proves } P_w(w)].$$

The statement in the square bracket is given partly in words, but it is a perfectly precisely defined statement. It asserts that the x-th proof is actually a proof of that proposition which is $P_w(\ )$ applied to the value w itself. In fact, because of the translations into arithmetic that we are supposing have been carried out, it is actually some *arithmetical* statement concerning the two natural numbers x and w. As before, precisely *which* arithmetical statement it is will depend on the details of the numbering systems, and it will depend very much on the detailed structure of the axioms and rules of our formal system. Since all that belongs to the complicated part of the argument, the details of it will not concern us here. Outside the square bracket the existential quantifier ($\exists$) serves to remove one of the variables, so we end up with an arithmetical propositional function which depends on only the one variable w. We have numbered all propositional functions which depend on a single variable, so the one we have just written down must have been assigned a number. Let us suppose that this number is k.

Then we have

$$\sim \exists x[\Pi_x \text{ proves } P_w(w)] = P_k(w).$$

Now examine this for the particular w-value: $w = k$. We get

$$\sim \exists x[\Pi_x \text{ proves } P_k(k)] = P_k(k).$$

The specific proposition $P_k(k)$ is a perfectly well-defined arithmetical statement. Does it have a proof within our formal system? Does its negation $\sim P_k(k)$ have a proof? The answer to both of these questions must be "no". We can see this by examining the *meaning* underlying the Gödel procedure. Although $P_k(k)$ is just an arithmetical proposition, we have constructed it so that it asserts what has been written on the left-hand side: "there is no proof, within the system, of the proposition $P_k(k)$". If we have been careful in laying down our axioms and rules of procedure, and assuming that we have done our numbering right, then it cannot be possible to find a proof of this $P_k(k)$ within the system. For if it were, then the "meaning" of the statement that $P_k(k)$ actually asserts would be false, so $P_k(k)$ would have to be false as an arithmetical proposition. Our formal system should not be so badly constructed that it actually allows false propositions to be proved! Thus, it must be the case that there is in fact *no* proof of $P_k(k)$. But this is precisely what $P_k(k)$ is trying to tell us. What $P_k(k)$ asserts must therefore be a *true* statement, so $P_k(k)$ must be true as an arithmetical proposition. It follows that we had also better not be able to find a proof of its negation $\sim P_k(k)$. We are not supposed to be able to prove false propositions within our formal system, and we have now seen that $\sim P_k(k)$ is certainly false! Thus, neither $P_k(k)$ nor $\sim P_k(k)$ is provable within our formal system. This establishes Gödel's theorem.

7. *Mathematical Insight*

Notice that something very remarkable has happened here. People often think of Gödel's theorem as something negative—showing the necessary limitations of formalized mathematical reasoning. No matter how comprehensive we think we have been, there will always be some propositions which escape the net. But should the particular proposition $P_k(k)$ worry us? As an essential ingredient of the above argument, we have actually established that $P_k(k)$ is a *true* statement! Somehow we have managed to *see* that $P_k(k)$ is true despite the fact that it is not formally provable within the

system. The strict mathematical formalists *should* indeed be worried, because by this very reasoning we have established that the formalist's notion of "truth" must be necessarily incomplete. *Whatever* formal system is used, there are statements that we can see are true but which do not get assigned the truth-value *"true"* by the formalist's proposed procedure, as described above. The way that a strict formalist might try to get around this would perhaps be not to talk about the concept of truth at all but merely refer to *provability* within some fixed formal system. However, this seems very limiting. (One could not even frame the particular form of the Gödel argument as given above, using that point of view, since parts of that argument made use of reasoning about what is actually true and what is not true.) Some formalists take a more "pragmatic" view, claiming not to be worried by statements such as $P_k(k)$ because they are extremely complicated and uninteresting as propositions of arithmetic. Such people would assert:

> "Yes, there is the odd statement, such as $P_k(k)$, for which my notion of provability or *"truth"* does not coincide with your instinctive notion of truth, but those statements will never come up in serious mathematics (at least not in the kind I am interested in) because such statements are absurdly complicated and unnatural as mathematics."

It is indeed the case that the statement $P_k(k)$ would be an extremely cumbersome and odd-looking mathematical proposition when written out in full. But there are closely related types of Gödel statement for which this is not so (cf. Paris and Harrington *1977*). The above seems to me to be a very strange point of view to adopt for a philosophy of mathematics. Moreover, it is not really all that pragmatic. When mathematicians carry out their forms of reasoning, they do not want to have to be continually checking to see whether or not their arguments can be formulated in terms of the axioms and rules of procedure of some formal system. They only need to be sure that their arguments are valid ways of ascertaining truth. The Gödel argument is another such valid procedure, so it seems to me that $P_k(k)$ is just as good a mathematical truth as any that can be obtained more conventionally using the axioms and rules of procedure that can be laid down beforehand.

Perhaps there are numerous perfectly acceptable results in the mathematical literature whose proofs require insights equally far from the rules and axioms of the original standard formal systems for arithmetic. Mathematical insights—by which I mean the means by which the mathematician arrives at his or her judgements of truth—are not something amenable to the procedures of formalization, and it would seem, therefore, that they are not the results of algorithmic operation.

Be that as it may, it is a clear consequence of the Gödel argument that mathematical truth is *not* an algorithmic matter; and it is something that goes beyond mere formalism. This is perhaps clear even without Gödel's

theorem. For how are we to decide what axioms or rules of procedure to adopt in any case when trying to set up a formal system? Our guide in deciding on the rules to adopt must always be our intuitive understanding of what is "self-evidently true", given the "meanings" of the symbols of the system. How are we to decide which formal systems are sensible ones to adopt—i.e., in accordance with our intuitive feelings about "self-evidence" and "meaning"—and which are not? The notion of self-consistency of the system is certainly not adequate for this. One can have many self-consistent systems which are not "sensible" in this sense. "Self-evidence" and "meaning" are concepts which would still be needed, even without Gödel's theorem.

However, without Gödel's theorem it might have been possible to imagine that the intuitive notions of "self-evidence" and "meaning" could have been employed just once and for all, merely to set up the formal system in the first place, and thereafter dispensed with as part of clear mathematical argument for determining truth. Perhaps these "vague" intuitive notions could have roles to play as part of the mathematician's preliminary thinking, as a guide towards finding the appropriate formal argument; but they would play no part in the actual demonstration of mathematical truth. Gödel's theorem shows that this point of view is not really a tenable one in a fundamental philosophy of mathematics. The notion of mathematical truth goes beyond the whole concept of formalism. There is something absolute and "God-given" about mathematical truth. This is what mathematical Platonism is about. Any particular formal system has a provisional and "man-made" quality about it. Such systems indeed have very valuable roles to play in mathematical discussions, but they can supply only a partial (or approximate) guide to truth. Real mathematical truth goes beyond these man-made constructions.

8. *Some Examples of Nonalgorithmic Mathematics*

There are very many areas of mathematics where problems arise which are not algorithmic. Thus, we may be presented with a class of problems to which the answer for each individual problem is either "yes" or "no", but for which no general algorithm exists for deciding which of these two answers is actually valid in each case. Some of these classes of problem are remarkably simple-looking.

First, consider the problem of finding integer solutions of systems of algebraic equations with integer coefficients. Such a set of equations might be

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

$$z^3 - y - 1 = 0, \quad yz^2 - 2x - 2 = 0, \quad y^2 - 2xz + z + 1 = 0$$

and the problem is to decide whether or not they can be solved for *integer* values of x, y, and z. In fact, in this particular case they can, a solution being given by

$$x = 13, \quad y = 7, \quad z = 2.$$

However, there is no algorithm for deciding this question for an arbitrary set of Diophantine equations. Diophantine arithmetic, despite the elementary nature of its ingredients, is part of nonalgorithmic mathematics!

As a somewhat less elementary-looking example, consider the question of topological equivalence of manifolds. Now, for ordinary surfaces—of *two* dimensions (say, closed and orientable)—there is, in fact, an algorithm for deciding whether or not two of them are topologically identical. This amounts, in effect, to counting the number of "handles" that each surface has. If the number of handles is the same in each case, then the two surfaces are topologically the same. Otherwise the surfaces are topologically different.

Now, suppose that we ask the same question for "surfaces" of *three* dimensions. Disconcertingly, the answer to this question is not known (at the time of writing). Let us pass on to the case of manifolds of *four* dimensions: is there an algorithm for 4-manifolds? Now the answer *is* known! There is *no* such algorithm. There is no algorithm for 5-manifolds either, nor for 6-manifolds, nor for any higher dimension. The question on the case of 4-manifolds is conceivably of some relevance to physics, since according to Einstein's general relativity, space and time together constitute a 4-manifold, and it has been suggested that this nonalgorithmic property might have relevance to "quantum gravity" (Geroch and Hartle *1986*). The fact that there is no algorithm for deciding whether two 4-manifolds are the same or not entails that, in a certain sense, 4-manifolds cannot be completely systematically classified. There is an unending variety of different kinds of structure that can occur.

Let us consider a different type of problem—one which is rather simpler to appreciate, namely the *word problem* (for semigroups). Suppose that we have some alphabet of symbols, and we consider various strings of these symbols, referred to as *words.* The words need not in themselves have any

meaning, but we shall be given a certain (finite) list of "equalities" between them which we are allowed to use in order to derive further such "equalities". This is done by making substitutions of words from the initial list into other (normally longer) words which contain them as portions. Each such portion may be replaced by another portion which is deemed to be equal to it according to the list. The problem is then to decide, for some given pair of words, whether or not they are "equal" according to these rules.

As an example, we might have, for our initial list:

$$\begin{aligned} \text{EAT} &= \text{AT} \\ \text{ATE} &= \text{A} \\ \text{LATER} &= \text{LOW} \\ \text{PAN} &= \text{PILLOW} \\ \text{CARP} &= \text{ME} \end{aligned}$$

From these we can derive, for example,

$$\text{LAP} = \text{LEAP}$$

by use of successive substitutions from the second, the first, and again the second of the relations from the initial list:

$$\text{LAP} = \text{LATEP} = \text{LEATEP} = \text{LEAP}.$$

The problem now is, given some pair of words, can we get from one to the other simply using such substitutions? Can we, for example, get from CATERPILLAR to MAN, or, say, from CARPET to MEAT? The answer in the first case happens to be "yes", while in the second it is "no". When the answer is "yes", the normal way to show this would be simply to exhibit a string of equalities where each word is obtained from the preceding one by use of an allowed relation. Thus (indicating the letters about to be changed in bold type, and the letters which have just been changed in italics):

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

CATERPILLAR C*A*RPILLAR CARPILL***ATER*** CARPIL***LOW***
CAR***PAN*** *ME*AN ME***ATE***N M***AT***EN
M*A*N.

How can we tell that it is impossible to get from CARPET to MEAT by means of the allowed rules? For this, we need to think a little more, but it is not too hard to see, in a variety of different ways. The simplest appears to be the following: in every "equality" in our initial list, the number of A's plus the number of W's plus the number of M's is the same on each side. Thus the total number of A's, W's and M's cannot change throughout any succession of allowed substitutions. However, for CARPET this number is 1 whereas for MEAT it is 2. Consequently, there is no way of getting from CARPET to MEAT by allowed substitutions.

Notice that when the two words are "equal" we can show this simply by exhibiting an allowed formal string of symbols, using the rules that we had been given; whereas in the case where they are "unequal", we had to resort to arguments *about* the rules that we had been given. There is a clear algorithm that we can use to establish "equality" between words whenever the words *are* in fact "equal". All we need do is to make some appropriate "alphabetical" listing of all the possible sequences of words of given length—and then successively increasing the length, as above; and then strike from this list any such string for which there is a pair of consecutive words where the second does not follow from the first by an allowed rule. The remaining sequences will provide all the sought-for "equalities" between words. However, there is no such obvious algorithm, in general, for deciding when two words are *not* "equal", and we may have to resort to "intelligence" in order to establish that fact. (Intelligence, incidentally, is useful also for establishing the *existence* of an "equality", but in that case it is not *necessary*!)

In fact, for the particular list of five "equalities" that constitute the initial list in this particular case, it is not unduly difficult to provide an algorithm for ascertaining that two words are "unequal" when they are indeed "unequal". However, in order to *find* the algorithm that works in this case we should have to exercise a certain amount of intelligence! For it turns out that there is no general algorithm which can be used universally for *all* possible choices of initial list. In this sense there is no algorithm for solving the word problem. The general word problem belongs to nonalgorithmic mathematics!

There are even certain *particular* selections of initial list for which there is no algorithm for deciding when two words are unequal. One such is given by

$$\begin{aligned} \mathrm{AH} &= \mathrm{HA} \\ \mathrm{HE} &= \mathrm{EH} \\ \mathrm{AT} &= \mathrm{TA} \\ \mathrm{TE} &= \mathrm{ET} \\ \mathrm{TAI} &= \mathrm{IT} \\ \mathrm{HEI} &= \mathrm{IH} \\ \mathrm{THAT} &= \mathrm{ITHT} \end{aligned}$$

(This list is adapted from one given by Tseitin and Scott *1955.*) Thus this *particular* word problem belongs to nonalgorithmic mathematics—in the sense that using this particular initial list we cannot algorithmically decide whether or not two given words are "equal". This type of word problem arose from considerations of formalized mathematical logic ("formal systems", etc.). The initial list plays the role of an axiom system and the substitution rule for words, the role of the formal rules or procedure. The proof that the word problem is nonalgorithmic arises from such considerations.

As a final example of a problem in mathematics which is nonalgorithmic, let us consider the question of covering the Euclidean plane with polygonal shapes, where we are given a finite number of different such shapes and we ask whether it is possible to cover the plane completely, without gaps or overlaps, just using these shapes and no others. Such an arrangement of shapes is called a *tiling* of the plane (Grünbaum and Shephard *1986*). We are all familiar with the fact that such tilings are possible using just squares, or just equilateral triangles, or just regular hexagons. There are many other single shapes which will tile the plane, some of which are quite elaborate. With a *pair* of shapes, the tilings can become more elaborate still. In fact, various pairs of shapes are known which will tile the plane only nonperiodically. For a *general* set of tiles there is no decision procedure (Berger *1966,* cf. also Robinson *1971*): the tiling problem, *also,* is part of nonalgorithmic mathematics! It is perhaps remarkable that such an apparently "trivial" area of mathematics—namely covering the plane with congruent shapes—which seems almost like "child's play" should in fact be part of nonalgorithmic mathematics. (The problem, as I have stated it, is actually somewhat more involved than the one treated by Berger. I am allowing polygons of general shape, whereas he used tiles based in squares, without rotations of the tiles. With the more general shapes, one would need some adequately computable way of displaying the individual tiles. One way of doing this would be to

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

give their vertices as points in the Argand plane given as algebraic numbers. The issues raised by these distinctions are not important for us here.)

9. *The Relevance of Complexity Theory*

The arguments that I have given above, and in the preceding chapters, concerning the nature, existence, and limitations of algorithms have been very much at the "in principle" level. I have not discussed the question of whether any algorithms which arise are likely to be in any way practical. Even for problems where it is clear that algorithms exist and how such algorithms can be constructed, it may require much ingenuity and hard work to develop the algorithms into something usable. Sometimes a little insight and ingenuity will lead to considerable reductions in the complication of an algorithm and sometimes to absolutely enormous shortening of its length of operation. These questions are often very detailed and technical, and it would not be pertinent for me to attempt to enter into a detailed discussion of them. There are various general things that are known, or conjectured, concerning certain *absolute* limitations on how much the speed of an algorithm can be increased. It turns out that even among mathematical problems that *are* algorithmic in nature, there are some classes of problems that are intrinsically vastly more difficult to solve algorithmically than others. The difficult ones can be solved only by very slow algorithms (or perhaps, with algorithms which require an inordinately large amount of storage space, etc.). The theory which is concerned with questions of this kind is called *complexity theory.*

Complexity theory is concerned not so much with the difficulty of solving *single* problems algorithmically but with infinite classes of problems where there is a general algorithm for finding answers to all the problems of the class. The different problems within such a class would have different "sizes", where the size of a problem is measured by some natural number n. The length of time—or more correctly, the number of elementary steps—that the algorithm would need for each particular problem of the class would be some natural number N which depends on n. To be a little more precise, let us say that among *all* of the problems of some particular size n, the greatest number of steps that the algorithm takes is N. Now, as n gets larger and larger, the number N is likely to get larger and larger too. Complexity theory has to do with the *rate of growth* of N, as n increases.

These questions are important for our considerations because they raise another issue, somewhat separate from that of whether or not things are algorithmic: namely, whether or not things that are known to be algorithmic are actually algorithmic in a *useful* way. I am inclined to think (though,

no doubt, on quite inadequate grounds) that, unlike the basic question of computability itself, the issues of complexity theory are not quite the central ones in relation to mental phenomena. Moreover, I feel that the questions of practicality of algorithms are being only barely touched by complexity theory as it stands today.

However, I could well be wrong about the role of complexity. Indeed, the complexity theory for *actual physical objects* could perhaps be different in significant ways from the standard complexity theory of algorithms. For such a possible difference to become manifest, it seems that it would be necessary to harness some of the remarkable properties of *quantum mechanics.* According to a recent set of ideas introduced by David Deutsch (*1985*), it seems to be possible *in principle* to construct a "quantum computer" for which there are (classes of) problems which could be solved by that device in polynomial time (i.e., with N increasing not faster than some polynomial in n) yet not by any Turing machine in polynomial time. It is not all clear, as of now, how an actual physical device could be constructed which behaves (reliably) as a quantum computer—and, moreover, the particular class of problem so far considered is decidedly artificial—but the *theoretical* possibility that a quantum physical device may be able to improve on a Turing machine seems now to be definitely with us.

10. Quantum Brains?

Can it be that a human brain, which I am taking for this discussion to be a "physical device"—albeit one of amazing subtlety, and delicacy of design, as well as of complication—is itself taking advantage of the magic of quantum theory? Do we yet understand the ways in which quantum effects might be used for benefit in the solving of problems or the forming of judgements? Is it conceivable that we might have to go even "beyond" present-day quantum theory to make use of such possible advantages? Is it really likely that actual physical quantum devices might be able to improve on the complexity theory for Turing machines?

What about the *computability* theory for actual physical devices? According to Deutsch's analysis, the class of computable operations is the *same* for a quantum computer as it is for a Turing machine. Consequently, one would need to go beyond standard quantum theory in order to be able to describe a device that could operate in a noncomputable way. This seems to be a tall order. Most physicists and physiologists would claim that there is no evidence even that ordinary quantum effects could be of any significance in the operation of a human brain, let alone putative effects that might lie outside standard quantum theory.

 On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

To address these matters I shall turn to a philosophical issue that I believe lends some support for quantum theory being relevant to our feelings of conscious awareness. Afterwards, I shall indicate why I believe that one must, indeed, go *beyond* standard quantum theory for a full picture.

An idea sometimes discussed in relation to conscious perception is the *teleportation machine* of science fiction (cf. Hofstadter and Dennett *1981*). This is intended as a means of "transportation" from, say, one planet to another—but whether it actually would be such is what the discussion is all about. Instead of being physically transported by a spaceship in the "normal" way, the would-be traveler is scanned from head to toe, the accurate location and complete specification of every atom in his body being recorded in full detail. All this information is then beamed (at the speed of light) by an electromagnetic signal to the distant planet of intended destination. There the information is collected and used as the instructions to assemble a precise duplicate of the traveler—together with all his memories, his intentions, his hopes, and his deepest feelings. At least that is what is expected; for every detail of the state of his brain has been faithfully recorded, transmitted, and reconstructed. Assuming that the mechanism has worked, the original copy of the traveler can be "safely" destroyed. Of course the question is: is this *really* a method of traveling from one place to another or is it merely the construction of a duplicate, together with the murder of the original? Would *you* be prepared to use this method of "travel"—assuming that the method had been shown to be completely reliable, within its terms of reference? If teleportation is *not* traveling, then what is the difference *in principle* between it and just walking from one room into another? In the latter case, are not one's atoms of one moment simply providing the information for the locations of the atoms of the next moment? We have seen, after all, that there is no significance in preserving the identity of any particular atom. The question of the identity of any particular atom is not even meaningful. Does not any moving pattern of atoms simply constitute a kind of wave of information propagating from one place to another? Where is the essential difference between the propagations of waves which describes our traveler ambling in a commonplace way from one room to another and that which takes place in the teleportation device?

Suppose it is true that teleportation *does* actually "work", in the sense that the traveler's own "awareness" is actually reawakened in the copy of himself on the distant planet (assuming that this question has genuine meaning). What would happen if the *original* copy of the traveler were not, after all, destroyed as the rules of this game demand? Would his "awareness" be in two places at once? (Try to imagine your response to being told the following: "Oh dear, so the drug we gave you before placing you in the Teleporter has worn off prematurely, has it, sir? That is a little unfortunate,

but no matter. Anyway, you will be pleased to hear that the other you—er, I mean the *actual* you, that is—has now arrived safely on Venus, so we can, er, dispose of you here—er, I mean of the *redundant* copy here. It will, of course, be quite painless.") The situation has the air of paradox about it. Is there anything in the laws of physics which could render teleportation *in principle* impossible? Perhaps, on the other hand, there is nothing in principle against transmitting a person—and a person's consciousness—by such means, but the "copying" process involved inevitably destroys the original? Might it then be that the preserving of *two* viable copies is what is in principle impossible?

I believe that despite the outlandish nature of these considerations, there *is* perhaps something of significance concerning the physical nature of consciousness and individuality to be gained from them. I believe that they provide one pointer, indicating a certain essential role for *quantum mechanics* in the understanding of mental phenomena. For it is a feature of standard quantum mechanics that a quantum state cannot simply be copied—unless the original state is destroyed in the process. If a mental state is like a quantum state, then the "teleportation paradox" may perhaps be resolved in this way (cf. Penrose *1987*).

Let us see how the point of view of strong AI relates to the teleportation question. We shall suppose that somewhere between the two planets is a relay station, where the information is temporarily stored before being retransmitted to its final destination. For convenience, this information is not stored in human form, but in some magnetic or electronic device. Would the traveler's "awareness" be present in association with this device? The supporters of strong AI would have us believe that this must be so. After all, they say, any question that we might choose to put to the traveler could in principle be answered by the device, by "merely" having a simulation set up for the appropriate activity of his brain. The device would contain all the necessary information; and the rest would just be a matter of computation. Since the device would reply to questions just exactly as though it were the traveler, then (Turing test!) it would *be* the traveler. This all comes back to the strong-AI assertion that the actual hardware is not important with regard to mental phenomena. This assertion seems to me to be unjustified. It is based on the presumption that the brain (or the mind) is, indeed, a digital computer. It assumes that no specific physical phenomena are being called upon, when one actually thinks, which might demand the particular physical (biological, chemical) structure that brains actually have.

No doubt it would be argued (from the strong-AI point of view) that the only assumption that is really being made is that the effects of any specific physical phenomena which need to be called upon can always be accurately *modeled* by digital calculations.

I feel fairly sure that most physicists would argue that such an assumption is actually a very natural one to make on the basis of our present physical understandings. However, it is precisely this assumption that I am disputing. If we accept this (commonly held) view that all the relevant physics *can* always be modeled by digital calculations, then the only real assumption (apart from questions of time, and calculation space) is the "operational" one that if something *acts* entirely like a consciously aware entity, then one must also maintain that it "*feels*" itself to be that entity.

11. Does Physics Allow Nonmechanisms?

But if our conscious thinking involves nonalgorithmic physical action, there would need to be scope for nonalgorithmic behavior in physical laws. This is assuming that—as I indeed believe—the actions of our brains or minds are indeed to be understood in terms of physical laws. What evidence is there that there is anything essentially nonalgorithmic that can be harnessed from the laws of physics? In discussing this question, it is necessary to make clear which laws one is regarding as relevant. We must ask what physics is likely to be actually involved in mental activity. Is Newtonian physics sufficient? Can we adequately regard our brains as functioning according to the laws governing the collisions and attractions between "Newtonian billiard balls"? Is there anything essentially new involved when we include the fields of Maxwell's electrodynamics? Does Einstein's special relativity have any significant role to play, or do the absurdly tiny space-time curvatures that arise because of the effects of his *general* relativity? What about quantum effects? Is there any conceivable reason to believe that quantum theory has relevance in the functioning of such a large and "warm" object as a human brain?

Even if so, do these more subtle physical theories provide any useful scope for nonalgorithmic action? Though there is little which is actually proved in this direction, I believe that there actually some good indication that there is *not* such scope in any useful sense. There is some interesting work which is related to these problems, by Pour-el and Richards (*1979, 1981, 1982*), which perhaps does point to the possibility of something nonalgorithmic, but I do not regard this as *useful* scope, if it is really scope at all.

Related to this is the kind of nonalgorithmic behavior that in effect arises when the gross performance of a system depends very delicately on its initial state. There are many situations when this is the case, such as in those circumstances which lead to "chaotic behavior". The large-scale behavior of a system may not be computable because it is not possible to know the initial

state to anything like the precision required. This kind of nonalgorithmic behavior, as it arises in practice, is certainly not what I would mean by "useful" nonalgorithmic action. It is not something that could be harnessed by the brain in order to achieve anything that a Turing machine cannot achieve. If we require something nonalgorithmic in the laws of physics that the brain might indeed be harnessing then, it seems to me, this must go *beyond* the physical laws that we now understand!

However, it is already hard to see how the physics of relativity or quantum theory could have any relevance to the functioning of the brain. It seems an even *more* remote possibility that there could be relevance for some *hitherto undiscovered theory* that goes beyond quantum theory and our now-standard picture of space-time according to relativity? Perhaps surprisingly, I think that there is some reason to believe that such an undiscovered theory *is* playing a role—indeed, even a functional role—in the actions of our conscious thinking. I cannot elaborate greatly on what I have in mind here, but roughly my picture is based on the following.

Quantum theory has been marvelously well substantiated at the molecular level and smaller. Here, the term "level" does not refer to spatial size but to the fact that only comparatively small amounts of mass-energy are involved. The effects of quantum theory are certainly maintained over substantial distances—at least over several meters as, in particular, the experiments of Aspect and co-workers (*1976, 1986*) strikingly demonstrate. However, the rules of the theory run into profound theoretical difficulties if applied to physical objects that are very much "larger" (in the sense of much larger mass-energy distributions, not physical size) than the levels at which the theory has been experimentally verified. Quantum theory insists that all the different alternative possible ways that a system *might* behave have to be allowed to *coexist* in (linear) superposition. At some stage, however, just *one* alternative out of this array of superposed alternatives finds realization in the actuality of experience. How does this happen? Quantum theory has no adequate explanation; and the question is debated endlessly by theoreticians. My own view (which is shared by a number of respectable people, I might add, including several of the founders of the subject!) is that quantum theory is a *provisional* theory, at the macroscopic level, and must eventually be replaced by something more satisfactory (and presumably nonlinear). I believe that persuasive reasons can be provided that these changes must come about at the stage when the differing mass-energy distributions that are to be linearly superposed have quantum-mechanically significantly differing *gravitational effects* (thus affecting the space-time geometry differently). Preliminary calculations indicate that such effects will already have become relevant at the level of one ten-millionth of a gram—but, I hope, sometimes also at significantly smaller levels (cf. Penrose *1986, 1987a*).

Can this sort of thing be relevant for the workings of the brain? I think that it is likely. It is already known that some nerve cells are sensitive to a single quantum stimulus (e.g., in a frog's retina, where a single photon can trigger a macroscopic nerve signal). For such cells, the effect is to convert something which has apparently to be treated quantum mechanically into something which has apparently to be treated classically. As yet, there is no adequate theory for this. My guess would be that for some classes of cells, or perhaps of chemical transmitters, in the brain, this kind of quantum/classical to-and-fro is an essential ingredient of our thinking. If this is the case, we would not, as yet, have an adequate theory to cope with it. The normal way that physicists would try to handle the situation would be to keep invoking the probabilistic rules which quantum theory uses whenever an "observation" is deemed to have taken place. But, in my opinion, these are only approximations, valid at the extreme limits when things go from the completely quantum to the completely classical (as with the standard "Copenhagen interpretation"). Here, I am imagining that things are more delicately poised at the quantum/classical borderline, where we really have no adequate theory at all.

I am guessing that the theory we shall need is essentially nonalgorithmic in nature. I believe that there are clear indications that it must also be nonlocal: that is to say, the total behavior cannot be considered to be composed of separated independent parts, but is, in some appropriate sense, "holistic". This kind of holistic behavior is already an observed feature of standard quantum theory, and is what the experiments of the Aspect type demonstrate. As yet, there is no indication from *physical* theory (as opposed to philosophical/mathematical considerations) that we shall need nonalgorithmic laws. But, as I indicated above, we do not yet have the correct physical theory. It should be made clear that the issue of nonalgorithmic behavior is quite a separate one from determinism or mathematical well-definedness. (For example, one can easily make up deterministic, though quite artificial, "toy models" which are not algorithmic. The state of our "toy physical system" could be given by a list of pairs of words. If the two words in the final pair are "equal" by virtue of all the other pairs being deemed "equal" then the list is to be modified in some preassigned way at the "next instant of time"; but if they are "unequal" the list is to be modified in a *different* preassigned way at the "next instant".) What I would mean by a physical "nonmechanism" would be some physical object whose behavior is governed in some such well-defined, but nonalgorithmic way.

Whether or not our physical theory eventually actually moves in such a direction, it has to be admitted that nonalgorithmic behavior is a possibility that must be seriously considered. It seems to me that the philosophical issues that this raises have barely begun to be explored. Alan Turing's sem-

inal work of fifty years ago may now be telling us that these possibilities are things that we may need to contend with in our understanding of the physical world. Perhaps, eventually, such considerations may help us to understand how it is that we can actually *be* part of that world!

References

Aspect, A.

1976 Proposed experiment to test the nonseparability of quantum mechanics. *Phys. Rev.* **D14** (1976) 1944- 51.

Aspect, A., and P. Grangier

1986 Experiments on Einstein-Podolsky-Rosen-type correlations with pairs of visible photons. In: *Quantum Concepts in Space and Time,* eds. R. Penrose and C.J. Isham. Oxford: Oxford University Press (1986).

Benacerraf, P.

1967 God, the Devil and Gödel. *The Monist* **51** (1967) 9–32.

Berger, R.

1966 The undecidability of the domino problem. *Mem. Am. Math.* **66** (1966).

Bowie, G.L.

1982 Lucas' number is finally up. *J. Philos. Lo.* **11** (1982) 279–285.

Deutsch, D.

1985 Quantum theory, the Church-Turing principle and the universal quantum computer. *P. Roy. Soc. A* **400** (1985) 97–117.

Geroch, R., and J.B. Hartle

1986 Computability and physical theories. *Found. Phys.* **16** (1986) 533.

Gödel, K.

1931 Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monats. Math. Phys.* **38** (1931) 173–198.

Good, I.J.

1969 Gödel's theorem is a red herring. *Br. J. Phil. S.* **18** (1969) 359–373.

Grünbaum, B., and G.C. Shephard

1986 *Tilings and Patterns.* New York: W.H. Freeman and Co. (1986).

Hofstadter, D.R.

1979 *Gödel, Escher, Bach: an Eternal Golden Braid.* Stanford Terrace, Hassocks, Sussex: the Harvester Press, Ltd. (1979).

1981 A conversation with Einstein's brain. In: Hofstadter and Dennett *1981.*

Hofstadter, D.R., and D.C. Dennett

1981 *The Mind's I.* Harmondsworth, Middx.: Basic Books, Inc., Penguin Books, Ltd. (1981).

Lucas, J.R.

1961 Minds, machines and Gödel. *Philosophy* **36** (1961) 120–124; reprinted in: Alan Ross Anderson, *Minds and Machines,* Englewood Cliffs (1964).

Mandelbrot, B.B.

1986 Fractals and the rebirth of iteration theory. In: *The Beauty of Fractals: Images of Complex Dynamical Systems,* eds. H.-O. Peitgen and P.H. Richter. Berlin: Springer-Verlag (1986) 151–160.

Paris, J., and L. Harrington

1977 A mathematical incompleteness in Peano arithmetic. In: *Handbook of Mathematical Logic,* ed. J. Barwise. Amsterdam: North-Holland Publ. Co. (1977).

Penrose, R.

1986 Gravity and state-vector reduction. In: *Quantum Concepts in Space and Time,* eds. R. Penrose and C.J. Isham. Oxford: Oxford University Press (1986) 126– 146.

1987 Minds, machines and mathematics. In: *Mindwaves,* eds. C. Blakemore and S. Greenfield. Oxford: Blackwells Publ. (1987).

1987a Newton, quantum theory and reality. In: *Three-Hundred Years of Gravity,* eds. S.W. Hawking and W. Israel. Cambridge: Cambridge University Press (1987).

Pour-El, M.B., and I. Richards

1979 A computable ordinary differential equation which possesses no computable solution. *Ann. Math. Log.* **17** (1979) 61–90.

1981 The wave equation with computable initial data such that its unique solution is not computable. *Adv. Math.* **39** (1981) 215–239.

1982 Noncomputability in models of physical phenomena. *Int. J. Theor.* **21** 553–555.

Robinson, R.M.

1971 Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.* **12** (1971) 177–209.

Searle, J.R.

1980 Minds, brains, and programs. In: *The Behavioral and Brain Sciences* vol. 3. Cambridge: Cambridge University Press (1980). Reprinted in Hofstadter and Dennett *1981.*

Tseitin, G.S., and D. Scott

1955 see *Logic Machines and Diagrams,* by M. Gardner, p. 144. Chicago: University of Chicago Press (1958).

On the physics and mathematics of thought. *The universal Turing machine: A half-century survey,* 491–522, 1988. Also published in: *Logic, Philosophy of Science and Epistomology* Proc. 11th Int. Wittgenstein Symp., 122–8, 1994/95 and 1998

Turing, A.M.

1936-7 On computable numbers, with an application to the Entscheidungsproblem. *P. Lond. Math. Soc. (2)* **42** (1936-7) 230–265; A correction, ibid. **43** (1937) 544–546.

1950 Computing machinery and intelligence. *Mind* **59** (1950) no. 236. Reprinted in Hofstadter and Dennett *1981.*

186
A TWISTOR CONFORMAL FIELD THEORY FOR FOUR SPACE-TIME DIMENSIONS

Phys. Lett. B **216** (1989), no. 1–2, 48–52. With M. A. Singer and A. P. Hodges

See Chapter 178.

786 A twistor conformal field theory for four space-time dimensions. *Phys. Lett. B* **216** (1989), no. 1–2, 48–52. With M. A. Singer and A. P. Hodges

A TWISTOR CONFORMAL FIELD THEORY FOR FOUR SPACE-TIME DIMENSIONS

A.P. HODGES, R. PENROSE and M.A. SINGER
Mathematical Institute, University of Oxford, 24–29 St. Giles, Oxford OX1 3LB, UK

Received 28 September 1988

A definition is proposed of a four-dimensional conformal field theory in which a class of complex three-manifolds and holomorphic sheaf cohomology replace the Riemann surfaces and holomorphic functions of two-dimensional theory. Suggestions are made about the interpretation of such a theory in terms of the interaction of fundamental particles, its relation to the theory of twistor diagrams, and the possibility of extending it to incorporate gravity.

In this note, we suggest a radical modification of the standard notion of two-dimensional conformal field theory (CFT2) – or of string theory – which provides the potential to represent the fundamental interactions of quantum fields in *four* space–time dimensions, while retaining the unifying complex-analytic ideas that have made CFT2 so attractive a model.

The essence of a CFT2 is that it yields a model of interacting quantum field theory, determined only by the intrinsic properties of a complex manifold rather than by equations or by action principles. Thus, given a Riemann surface X, whose boundary ∂X consist of n positively oriented and m negatively oriented parametrized circles, there is a rule which determines (in a natural way) a ray in the product of Hilbert spaces $\mathscr{H}^{\otimes n} \times \mathscr{H}^{*\otimes m}$, where $\mathscr{H}$ is the fermionic Fock space on the Hilbert space of positive-frequency complex-values functions (strictly half-forms) on S^1 and $\mathscr{H}^*$ is its dual, arising from the space of negative-frequency functions on S^1. The required ray is determined by considering, in effect, the *Grassmann representation* of the *linear space of holomorphic functions* $\mathscr{O}(X)$ on X, restricted to ∂X (and assumed smooth on ∂X), as a subspace of the space of all (C^∞) functions on X:

$$\mathscr{O}(X) \subset C^\infty(\partial X)$$

(for details, see ref. [1]). A crucial feature of this construction is the splitting property (essentially by Laurent series) whereby a smooth function on the circle, which we now take to be the equator $S = \{z: |z| = 1\}$ of the Riemann sphere, is the sum of a "positive frequency" holomorphic function, extending to the northern hemisphere $\{z: |z| < 1\}$, and a "negative frequency" holomorphic function extending to the "southern hemisphere" $\{z: |z| > 1$, or $z = \infty\}$.

Positive- (negative-) frequency data on the n bounding circles of a Riemann surface X, can be though of as analogous to in- (out-) states in a dynamical theory, so that knowledge of the ray defined by X is equivalent to knowing, up to an overall phase, all "S-matrix elements" between such states. Thus, CFT2 allows the Riemann surface X to determine the "interaction" of these states. An attractive feature of this picture is the hope that properties of surfaces of higher genus (with "handles") may embody the essentially quantum-field-theoretic content of the "loops" of Feynman diagrams. (However, for such an interpretation to hold, some rule must ultimately be given for summing over the different genera and moduli of Riemann surfaces X which interpolate between n circles.)

Our proposal is to provide a radical re-interpretation of this type of picture, whereby the Riemann surfaces of CFT2 are replaced by (generalized) *twistor spaces* – either three-complex-dimensional *projective* twistor spaces of their four-complex-dimensional non-projective counterparts. We consider just the projective case here, and then such a twistor space P would be a complex three-manifold, with boundary ∂P consisting of $n+m$ standard (parametrized) copies of a five-real-dimensional space $\mathbb{PN}$, n of which

are positively oriented and m negatively oriented (fig. 1). (Standard projective twistor space $\mathbb{PT}$ is a complex projective three-space $\mathbb{CP}^3$, and $\mathbb{PN}$ is the subspace of $\mathbb{PT}$ defined by the vanishing of a hermitian form Σ of signature $++--$; see refs. [2,3] for details.) The points of $\mathbb{PN}$ are interpreted as the *light rays* (null straight lines) in ordinary (compactified) Minkowski four-space $\mathbb{M}$.

In standard twistor theory, massless free fields (of arbitrary spin) are succinctly represented as elements of *first holomorphic sheaf cohomology* – i.e. by holomorphic one-functions – rather than by the ordinary holomorphic functions (zero-functions) of CFT2 (see refs. [2,3]). In place of the splitting of functions on the circle S into positive- and negative-frequency parts according to whether they extend into the northern or southern hemisphere, we now have a corresponding splitting of one-functions defined on $\mathbb{PN}$ into positive- and negative-frequency parts depending on whether they extend, as holomorphic one-functions, into the top half $\mathbb{PT}^+$ ($\Sigma>0$) or bottom half $\mathbb{PT}^-$ ($\Sigma<0$) of $\mathbb{PT}$. This splitting corresponds precisely to the physical requirements for the in- and out-states of massless fields.

A massless field of helicity s is described by a holomorphic one-function which is homogeneous of degree $-2s-2$, and we denote the space of such one-functions, defined on the complex three-manifold P, by $H^1(P, \mathcal{O}(-2s-2))$. This space plays the role of the $\mathcal{O}(X)$ of CFT2. We define a CFT4 in a way closely corresponding to CFT2, but with the inclusion

$$H^1(P, \mathcal{O}(-2s-2)) \subset H^1(\partial P, \mathcal{O}(-2s-2))$$

taking over the role of $\mathcal{O}(X) \subset C^\infty(\partial X)$. (The space on the right is defined in terms of "CR-cohomology", ∂P being not a complex manifold but a *boundary* of a complex manifold; see ref. [4]. For our purposes, we define a "CR-manifold" to be an odd-dimensional manifold, such as $\mathbb{PN}$, with intrinsic structure defined from its being a hypersurface in a complex manifold.) The roles of $\mathcal{H}$ and $\mathcal{H}^*$ are taken over by the fermionic Fock spaces (including anti-particles) on

$$H^1(\mathbb{PT}^+, \mathcal{O}(-2s-2)), \quad H^1(\mathbb{PT}^-, \mathcal{O}(-2s-2)),$$

respectively, and the Grassmann representation of $H^1(P, \mathcal{O}(-2s-2))$ serves to define our CFT4. This provides an interacting quantum field theory, but dependent (as with a CFT2) upon the particular choice of space P. One needs somehow to be able to integrate over the "space of possible P's" in order to obtain a theory which is invariant under the conformal group [effectively $SU(2,2)$] of $\mathbb{M}$.

Note that our generalization of CFT2 to four dimensions is completely different from what one would obtain by regarding the circle S as a bounding hypersurface in a two-dimensional space–time, and then generalizing to four space–time dimensions by considering a kind of higher-dimensional "membrane" theory (here "three-branes"). We do better than this; for by generalizing to twistor space, we may preserve

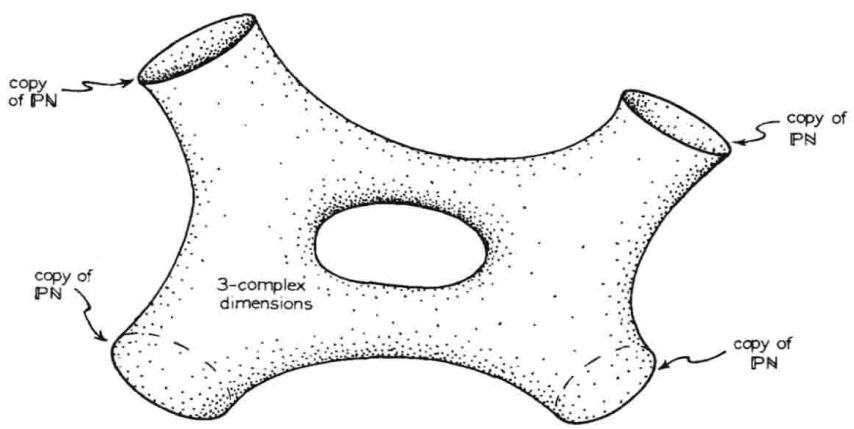


Fig. 1. The generalized twistor space P is a complex manifold of three dimensions, with a boundary consisting of disjoint copies of the five-real-dimensional space $\mathbb{PN}$. (Each point of $\mathbb{PN}$ represents a light ray in Minkowski space.)

not only the idea of a bounding hypersurface but also the key features of splitting and complex-analyticity which give the CFT2 its profound connection with quantum field theory.

Although we have stressed the close analogy between CFT2 and our proposed CFT4, there are notable differences. In two dimensions the notions of conformal manifold and complex manifold coincide, so that a conformal theory is naturally also a holomorphic theory; but in four dimensions they do not, and in describing CFT4 as a conformal field theory in four dimensions, we refer to the conformal geometry of $\mathbb{M}$, this being encoded in the *holomorphic* geometry of $\mathbb{PT}$. Conformal invariance in $\mathbb{M}$ is actually secondary to the holomorphic and "splitting" properties. (Indeed, as we indicate below, developments of our theory involve *breaking* conformal symmetry in $\mathbb{M}$, whilst retaining the unifying complex-analytic ideas.) There is another sharp difference between the cases of two and four dimensions, and the four-dimensional case is actually simpler. For the space of parameterizations of the circle is infinite-dimensional – giving the group $\mathrm{Diff}(S^1)$ – but the space of allowed parameterizations of $\mathbb{PN}$ is finite-dimensional, namely $SU(2, 2)$. (A parametrized copy of $\mathbb{PN}$ is a CR-manifold that is CR-equivalent to $\mathbb{PN}$, together with a choice of such equivalence.)

We must now justify our implication that there exist satisfactory analogues, in three complex dimensions, of the general (bounded) Riemann surfaces that give the CFT2 its content. At first sight, it might appear that one should regard any complex three-manifold as such an analogue. But the requirement that such a manifold should contain copies of $\mathbb{PN}$ (as a CR-manifold) imposes severe constraints. (This is in sharp contrast with CFT2, since any Riemann surface contains copies of S.)

This observation is naturally met by the class of *flat twistor spaces* P, defined according to

(i) each point of $P-\partial P$ possesses a neighbourhood biholomorphic to a neighbourhood of a line in $\mathbb{CP}^3$,

(ii) ∂P is CR-equivalent to a disjoint union of copies of $\mathbb{PN}$,

(iii) the canonical bundle of P admits a fourth root.

Flat twistor spaces are plentiful. For example, the twistor space of a conformally flat riemannian four-manifold M (cf. ref. [5]) always satisfies (i); it satisfies (iii) iff M has spin structure (cf. ref. [6] – we are grateful to Claude Lebrun for discussions on this point); and an appropriate boundary satisfying (ii) arises if the boundary of M consists of a number of disjoint (round) spheres. More generally we can take P to be the quotient of an open subset of $\mathbb{CP}^3$ by the free action of a group with g generators. Less formally such manifolds arise by the attachment of g "handles" to $\mathbb{CP}^3$ (much as the general Riemann surface arises by attaching handles to the Riemann sphere), and where there can also be h "holes" each of which is obtained similarly to the removal of $\mathbb{PT}^-$ from $\mathbb{PT}$. For fixed g and h there is a moduli space of (complex) dimension $15(g+h-1)$, except that the dimension is 0 if $g=0=h$ and 3 if $g=1$, $h=0$. (Compare this with the classical formula $3g-3$ for the number of moduli of a closed Riemann surface of genus $g>1$.)

In our favoured interpretation, the CFT4 picture is analogous to a Feynman diagram and so the (one-function) data on each copy of $\mathbb{PN}$ should represent a single incoming or outgoing particle. But as with CFT2, the ray in $\mathscr{H}^{\otimes n}\times\mathscr{H}^{*\otimes m}$ that the theory defines provides scattering amplitudes between different *Fock space* elements, one such element for each S^1 component of ∂X. Only at the lowest "one-particle" level of Fock space do we get something analogous to a Feynman diagram. How then do we interpret the "many-particle" levels of Fock space in our theory? An element of the nth level of the Fock space should be associated with a *single* holomorphic function (or appropriate cohomological generalization of a holomorphic function) of n twistors. Such functions of two, three or more twistors have long been studied within twistor theory. They have the important property of being able to represent particles with mass, and possess an internal symmetry structure of considerable interest [7,8]. (Conformal invariance is broken by the explicit picking out of the point at infinity in $\mathbb{M}$, in the specification of these functions.) Thus, although we defined our Hilbert space in terms of massless fields, there is the potential to include massive fields as well.

Suppose that (as is normal also in string theory) we restrict attention to the first, i.e. "one-particle", level of the Fock space. Then, a given flat twistor space P will define a (multilinear) functional of massless fields. Such a functional would have an interpretation as a scattering amplitude, and hence correspond

to some interaction of massless fields in $\mathbb{M}$. It would be necessary to sum over all admissible manifolds, of all genera and moduli, to obtain a meaningful amplitude, and it appears that some further substantial ingredient in our theory is required (as with CFT2) to make sense of this idea. But suppose this further ingredient has been supplied, then we naturally ask *what* interaction could turn out to be represented by the CFT4, and whether it could coincide with the interactions between massless fields which appear in the standard model.

In studying this possibility, we note that there already exists an approach within twistor theory to the description of quantum field-theoretic interaction, namely that of *twistor diagrams*. These are many-dimensional complex contour integrals which convolve together the twistor representatives of (massless) in- and out-fields with certain simple "propagator" kernels, yielding (multilinear) functionals corresponding to the scattering amplitudes defined by Feynman diagrams [9]. Recent progress has indicated that all Feynman diagrams for interactions appearing in the standard model (including QED) may be translated systematically into such a twistor integral form. (This involves passing to *non-projective* twistor space $\mathbb{T}$.) We can also include massive fields with many-twistor functions and explicit breaking of conformal invariance.

However, even if this programme of translation is to be entirely successful, it will result in a perturbative expansion with a generating principle borrowed from the standard treatment of the lagrangian in space–time, rather than defining a dynamical theory in its own right. The twistor diagram calculus also exhibits an unsatisfactory feature of non-uniqueness: many diagrams may yield the same quantum-field-theoretic amplitude. The situation is reminiscent of the dual diagrams which arose early in the theory of hadronic scattering. In that case the non-uniqueness could be elegantly resolved by identifying equivalent diagram sums as alternative evaluations of a scattering amplitude properly and uniquely defined on a (bosonic) string structure. It seems likely that the twistor diagram calculus analogously constitutes a contour integral formalism for the calculation of scattering invariants which are properly and uniquely defined in terms of some deeper geometric structure, the multiplicity of equivalent diagrams merely arising as different methods for evaluating the same quantity.

We are now led to suggest that this deeper structure might be just that of a CFT4: the diagrams would constitute an evaluative integral calculus for the theory, and the CFT4 theory itself would supply the missing generating principle for the diagrams.

In our final comments we refer to the question of incorporating gravitation into the picture, which must also involve a breaking of conformal invariance. This appears to be possible if we now allow $\mathbb{PT}$ to become deformed. The most clear-cut way of doing this, for our present purposes seems to be to take our twistor spaces to be the *asymptotic twistor spaces* [3,10] for asymptotically flat space–times. For every such space–time there would be an ingoing asymptotic projective twistor space $\mathbb{P}\mathcal{T}_{\text{in}}$ and an outgoing one $\mathbb{P}\mathcal{T}_{\text{out}}$ (each three-complex-dimensional). Each of these would contain a five-real-dimensional hypersurface ($\mathbb{P}\mathcal{N}_{\text{in}}$ and $\mathbb{P}\mathcal{N}_{\text{out}}$ respectively). The structure of $\mathbb{P}\mathcal{T}_{\text{in}}$ (or of $\mathbb{P}\mathcal{T}_{\text{out}}$) as a complex manifold codes the information of the anti-self-dual part of the incoming (or outgoing) gravitational radiation field [11]; and the location of $\mathbb{P}\mathcal{N}_{\text{in}}$ within $\mathbb{P}\mathcal{T}_{\text{in}}$ (or of $\mathbb{P}\mathcal{N}_{\text{out}}$ within $\mathbb{P}\mathcal{T}_{\text{out}}$) codes the self-dual part (cf. ref. [12]).

For a generalized twistor space $\mathcal{P}$, with many components to the boundary, there would be severe constraints on the different versions of $\mathbb{P}\mathcal{N}_{\text{in}}$ and $\mathbb{P}\mathcal{N}_{\text{out}}$ which could together constitute these various components. The very structure of $\mathcal{P}$ now partly takes over (i.e. for the gravitational field) the role that the holomorphic one-function had before, and further ideas are needed in order that our "Grassmann" construction can be carried over to this case. Such ideas would also have implications for how one is to "sum over" the different twistor manifolds that can arise in a CFT4.

References

[1] E. Witten, Quantum field theory, grassmannians, and algebraic curves, Commun. Math. Phys. 111 (1988), to be published;
G.B. Segal, The definition of conformal field theory, to be published, (1988).

[2] S.A. Huggett and K.P. Tod, An introduction to twistor theory (Cambridge U.P., Cambridge, 1985).

790 A twistor conformal field theory for four space-time dimensions. *Phys. Lett. B* **216** (1989), no. 1–2, 48–52. With M. A. Singer and A. P. Hodges

[3] R. Penrose and W. Rindler, Spinors and space-time, Vol. 2 (Cambridge U.P., Cambridge, 1986).
[4] R. Penrose, Bull. Amer. Math. Soc. 8 (1983) 427.
[5] M.F. Atiyah, N.J. Hitchin and I.M. Singer, Proc. R. Soc. London A 362 (1978) 425.
[6] N.J. Hitchin, Proc. London Math. Soc. (III Ser.) 43 (1981) 133.
[7] R. Penrose, Twistors and particles, in: Quantum theory and the structures of time and space, eds. L. Castell et al. (Carl Hanser Verlag, Munich, 1975).
[8] L.P. Hughston, Twistors and particles, Lecture Notes in Physics, Vol. 97 (Springer, Berlin, 1979).
[9] A.P. Hodges, Proc. R. Soc. London A 397 (1985) 341.
[10] R. Penrose and M.A.H. MacCallum, Phys. Rep. 6 (1972) 241.
[11] R. Penrose Gen. Rel. Grav. 7 (1976) 31.
[12] M. Ko, E.T. Newman and R. Penrose, J. Math. Phys. 18 (1977) 58.

187
MINDS, MACHINES AND MATHEMATICS

Mindwaves. Eds C. Blakemore and S. Greenfield. Blackwell, Oxford, 1989

See also Chapters 170, 185, and the summary, Chapter 195, of my book *The Emperor's New Mind* (Oxford University Press, 1989).

18

Minds, Machines and Mathematics

Roger Penrose

Roger Penrose was born at Colchester, Essex, in 1931, the second of four children of the geneticist Lionel S. Penrose and Margaret (née Leathes). He was supposed to become a doctor (as an antidote to all the chess-playing in the family – and his secret ambition had been to be a brain surgeon), but he liked puzzles and geometry and so became a mathematician instead. After stints at Cambridge, London (mainly Birkbeck College) and elsewhere, he became, in 1973, Rouse Ball Professor of Mathematics at Oxford (and, later, part-time Lovett Professor at Rice University, Houston, Texas). He has worked on black holes, pentagonal crystals, twistors and other aspects of mathematical physics. He has collected the odd prize, medal or FRS, but fortunately no honorary degrees. He has three sons but is divorced. Though he is considered an expert on space and time, he claims to be unable to organize either.

Now that the fruits of electronic computer technology are beginning to impinge on various aspects of our daily lives, it is becoming particularly topical to address the question of whether such computers can be said actually to think. To what extent are our own thought processes like the functioning of a computer? Are computers likely soon to acquire genuine intelligence? May they soon exceed our own abilities in this respect? Does it matter what a 'thinking device' is actually made of, or is its computational algorithm all that counts? Indeed, to what extent is it appropriate to analyse our own thought processes in terms of such computational algorithms? My purpose here is to make some remarks relevant to these matters.

Algorithms and Turing Machines

First, I should explain what is meant by a *computational algorithm* (or recursive process). Basically, this is an operation that can be performed by a computer – though one has to be a little bit careful because, strictly for the purpose of the definition, we mean a computer with an unlimited

In a remarkable series of experiments, D. A. Baylor, T. D. Lamb and K.-W. Yau demonstrated the electrical responses of individual rod photoreceptors from the eye of the toad in response to the absorption of single photons of light – the smallest indivisible units of light energy. This photograph shows a small fragment of toad retina with one of the rods sucked up into a glass tube containing fluid connected to a sensitive amplifier detecting any electric current flowing through the rod. A bright bar of light is shone on the glass tube, so that light falls on the rod and sets off an electrical response (photograph supplied by the authors, reprinted by permission of the Journal of Physiology*).*

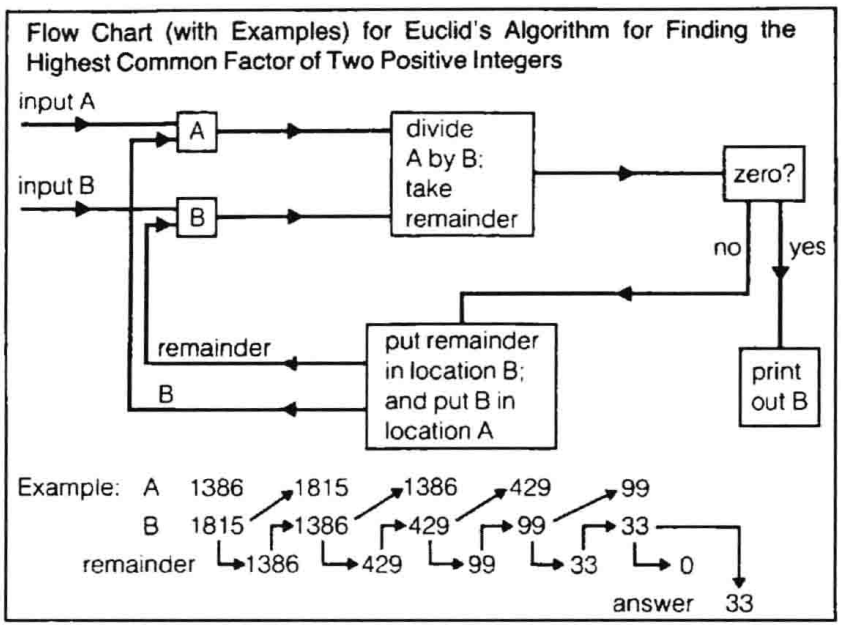
memory store which is allowed to continue for an unlimited time without wearing out. All this can be made precise in terms of the abstract mathematical object known as a *Turing machine* (which is the concept on which modern electronic computers are actually based).[1] In a Turing machine, any algorithmic process is broken down into elementary steps, where at each step the operation to be carried out is perfectly clear-cut and immediate, and so also is the procedure for moving on to the next elementary step. There is, moreover, a concept referred to as a *universal* Turing machine – which is some particular Turing machine with the capacity to imitate any other Turing machine. All one needs to do is to feed the universal machine first with a positive integer, say N, and it will then behave, in effect, as though it were the Nth Turing machine. Modern general-purpose computers are, in essence, universal Turing machines.

Algorithm

An entirely mechanical mathematical procedure, which is completely well defined in operation and which can be applied automatically without any additional decisions, judgements or random choices having to be made. The word comes from the name of a ninth-century Arab mathematician *al-Khowârizmi*. He was the author of a Persian textbook on algebra.

In our present century, the British mathematician, logician and code-breaker Alan Turing (1912–54) formalized the notion of algorithm in terms of the concepts which has become known as a (universal) Turing machine. This logical concept lies at the heart of all modern general-purpose computers.

The word 'algorithm' refers to a procedure for solving mathematical problems in a specified class, where there is a clear-cut rule to follow at each stage – that is to say, it is a procedure which can be carried out by a Turing machine. As an example, there is the well-known procedure referred to as *Euclid's algorithm*, which finds the highest common factor of two positive integers. Divide the smaller integer into the larger one; take the remainder; divide that into the previous smaller one; take the remainder again, and so on until the division goes exactly. The final (non-zero) remainder is the required highest common factor. It is very simple to write a flow diagram for this procedure (see box). We are familiar with many far more complicated (and extremely impressive) procedures that can now be carried out by computers. But these procedures are still composed of the same standard elementary clear-cut steps at each stage. (It can happen that sometimes such calculations simply go around in loops and do not terminate. Or perhaps they may just expand without limit. In such cases the calculations will not be as useful as an algorithm. For a proper algorithm we should be assured of an answer in a finite time in each case.)


Flow Chart (with Examples) for Euclid's Algorithm for Finding the Highest Common Factor of Two Positive Integers
input A
input B
A
B
divide A by B: take remainder
zero?
no
yes
remainder
B
put remainder in location B; and put B in location A
print out B
Example: A 1386 1815 1386 429 99
B 1815 1386 429 99 33
remainder 1386 429 99 33 0
answer 33

The Turing Test

I shall need to consider what is known as the Turing test.[2,3] This is intended to be an objective way of deciding whether a computer could validly be said to have been programmed to 'think'. The idea is that the computer and a human subject are both to be hidden from the view of a perceptive experimenter. The experimenter can communicate with either of the two subjects, human or electronic, but is not told which is which.

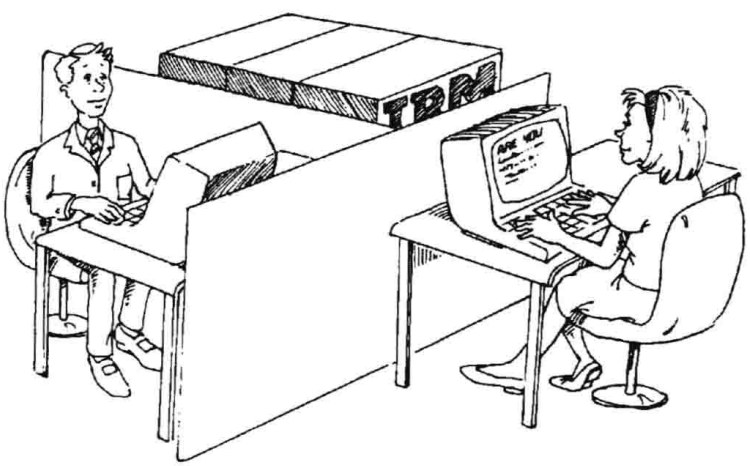


Figure 18.1 The Turing test: the experimenter, depicted at the right, has to decide which of two hidden subjects is a computer and which is a human being.

The questions that the experimenter puts and the answers that the subjects provide are each conveyed in an impersonal way – say, typed in and displayed on a screen. The experimenter, by such suitable questioning, is to decide which is the human subject and which is the electronic one. If the experimenter cannot reliably tell which is which, then the computer (or computer's programmer) has passed the test.

Of course, part of the skill involved in getting the computer to pass would be to make it *suppress* the kind of answers that a computer could easily provide, but which an ordinary human subject could not. If, for example, the experimenter were to ask some complicated arithmetical question, the computer would have to 'pretend' not to be able to answer it easily. And, of course, the computer would have deliberately to 'lie' sometimes – such as in replying to the direct question 'Are you a human being?' But all that would be the easy part. The hard part would be in getting the computer to give convincing answers to questions that could be answered easily in a 'common-sense' way by an ordinary average human subject.

The 'strong AI (artificial intelligence) point of view[3,4] asserts not only that it is merely a matter of (a comparatively short) time before a computer will be able to pass the Turing test, but also that once it has done so, terms such as

thought
understanding
awareness
happiness
pain
compassion
pride

could then be applied as validly to the computer as to the human subject. Such a viewpoint provides a very 'operational' approach to such questions. But strong-AI supporters would argue that the operational approach is necessary for cutting away the nebulous and shoddy thinking that pervades discussions in this kind of area – and it provides, instead, a rigorous and scientifically testable framework.

We recall that any universal Turing machine can, if supplied with an appropriate program, imitate any other Turing machine. Moreover, present-day general-purpose electronic computers are, in effect, already instances of universal Turing machines. Thus all that is now needed, from the point of view of *hardware*, for a modern computing machine, is a large enough store and a sufficient rapidity of operation that it can carry out its instructions adequately fast. Thus, so the strong-AI argument goes, it is merely the nature of the computer's program – that is, its software – which determines whether 'thinking', 'understanding', or any other of the terms in our list, has become appropriate to a modern computer's functioning.

Searle's Chinese Room

Not everyone accepts this view, however. In particular, John Searle in his

chapter in this book and elsewhere,[5] has argued strongly that 'understanding' cannot simply be a matter of the carrying out of an appropriate 'program' or 'algorithm'. He cites his example of the 'Chinese room'. This is a thought experiment based on some computer program (such as that of Roger Schank) which can correctly answer simple but not quite direct questions about certain elementary stories which Searle takes in Chinese. Searle envisages the program being enacted by the manipulation of Chinese symbols to obtain the desired answers. We may ask whether a computer following such a program can actually be said to 'understand' the story. The strong-AI answer would appear to be 'Yes'. But Searle says 'No'. For the electronics of the computer are replaced in Searle's thought experiment by a human worker – or perhaps by a room full of human workers, none of whom understands Chinese – each putting counters with Chinese symbols on them in various bins – and taking them out again –according to a set of instructions given to the workers in a book of rules printed, say, in English. The workers clearly do not understand the story since none of them understands Chinese. Still less does the room itself understand the story!

In my opinion, Searle has quite a strong case. But it is not a totally overwhelming one. The computer programs that do exist at present, and which are designed to provide answers to simple questions about such a restricted class of stories, are possibly just simple enough that one might reasonably envisage a person, or perhaps a room full of people carrying out the appropriate instructions. But when an actual Chinese person comprehends a Chinese story, the algorithm that is being carried out in that Chinese person's head – if, indeed, the followers of 'strong AI' are correct in believing that it *is* some algorithm – must be incomparably more complicated. The complication might be so great that imitation by a room full of human workers is simply out of the question. (It is not unknown in physics that totally new effects can arise, with 'big' systems, that are not discernible at all at a smaller scale.) Nevertheless, I feel that Searle is probably correct in maintaining that 'understanding' is not just a question of the carrying out of some appropriate algorithm – and still less is it merely a matter of the *output* of an algorithm. (Different algorithms can sometimes produce identical outputs.) It would seem, if this is the case, that 'understanding' is *not* just a question of *software* (that is, of programs or algorithms), but that the *hardware* should also be important. By 'hardware' is meant the actual physical construction of the object – either computer or brain.

I am not sure that Searle would put it quite like that – I am rather freely interpreting what I take to be his position (or the implication of his position) – but according to a viewpoint such as his, one would seemingly be driven to believing that a computer designed to follow the logical operation of every detail of the workings of a human brain would itself *not* achieve 'understanding' even though that *person* whose brain is under consideration may claim to be attaining understanding in some matter. The difference between this brain and this computer would not lie in the algorithms they are performing – since these are taken to be identical – but (presumably) only in their actual physical construction, that is, in their 'hardware'. (This contrasts markedly with the 'strong-AI' views expressed by, for example, Douglas Hofstadter.[4])

Non-algorithmic Functioning

My own viewpoint is not really the same as this. While I do indeed believe that the nature of the hardware is likely to be vitally important, I do not think that this can be separated from the question of software. I shall try to argue that if the 'hardware' is chosen appropriately, then its functioning is not something that is describable in terms of a computational algorithm at all. How this could come about I shall try to indicate shortly. For the moment, let us assume that non-algorithmic functioning *is* possible, and that a conscious human (or non-human) brain may well be non-algorithmic in its behaviour.

One of the implications of this viewpoint is that I do not really expect that an ordinary algorithmic computer could properly pass the Turing test at all. I should, however, at least qualify that statement. It could well be that a sufficiently clever program will some day be able to fool a perceptive experimenter for some of the time. It depends to some extent on how long the experimenter is allowed to continue the questioning in one single test and on how probing and sustained that questioning is allowed to be. Also, one could imagine a situation in which a computer is involved in a long succession of failures of the Turing test, but then the various answers provided by the human subject are 'studied', 'analysed' and 'remembered' by the computer and subsequently trotted out, perhaps with some random element, as the computer's 'own' answers at some later date. It is perhaps unfair to expect the experimenter to continue to ask original types of questions indefinitely and it is easy to imagine that after a long succession of separate tests the experimenter could get fooled. Somehow I regard this as 'cheating' and not 'properly' passing the test – though I admit the distinction is a bit vague. At least I doubt that many would regard *such* passing of the test as providing good evidence that the computer was actually 'thinking' or 'understanding' in the normal human way.

My own view is that if 'machines' are ever built that can *properly* (and consistently) pass the Turing test, then such 'machines' would have to incorporate the same non-algorithmic elements as are present (I am claiming) in conscious human thinking. More precisely what I mean by 'non-algorithmic elements' I shall have to explain later. For the moment let me just say that I do *not* mean simply *random* elements. I mean instead that some essentially non-algorithmic behaviour inherent in physical laws is being called upon.

The Active Role of Consciousness

In claiming that conscious thinking can invoke such non-algorithmic behaviour I am, in effect, tacitly claiming that consciousness is *functional*; that is, rather than being merely an 'accidental' feature of our particular (say, biological) construction – or, perhaps, as the strong-AI point of view might have it, being a necessary concomitant of sufficiently complex or high-level modes of computation ('thinking'?) – consciousness (I claim) has an *active* role to play in behaviour. Such a viewpoint seems to be rather unpopular with scientifically minded people (the *passive* role of conscious-

ness being readily admitted), perhaps because it seems to imply that this active role of consciousness goes 'outside science' – that is, that the laws of physics must be suspended in order to allow this conscious 'self' to affect the body's behaviour. However, I am making no such suggestion, as we shall see. The philosophical distinction between consciousness playing an active role, or merely a passive one, is perhaps a delicate one. I shall not concern myself with that question here. My use of the word 'active' in this context has mainly a motivational significance. I think it is a useful way of thinking about consciousness, particularly in connection with the question of how consciousness can have arisen as a result of natural selection.

I should make the (obvious) point that a good deal of the (very complicated) activity that goes on in our heads is not in fact conscious. A good deal of our behaviour is the immediate result of unconscious rather than conscious 'thinking'. When I say that consciousness is playing an active role I mean that the kind of behaviour which seems necessarily to be accompanied by conscious thought is different in character from that which does not. One of the things that makes this difficult to analyse, however, is that activities which at *one time* may have required consciousness may, later on, be carried out entirely unconsciously. Driving a car is one obvious example, or riding a bicycle, or (presumably) walking. Such activities are commonly carried out when one *is* conscious though not actually conscious of the activities themselves. Yet there is still a marked difference in the manner in which they are carried out, according to whether or not one's consciousness is actually brought to bear on those activities. This is clear from the manifestations of those circumstances when it is not so brought to bear, such as in day-dreaming. Moreover, consciousness may be absent altogether, such as in sleep-walking or (more pathologically) cases of epileptic 'fugue', in which a sequence of actions is automatically carried out. Here, the absence of conscious direction can be particularly manifest in the resulting behaviour.

Figure 18.2 In states such as sleepwalking, the effects of lack of conscious direction can be particularly manifest.

There is, moreover, a clear and perhaps rather trivial way in which the activities of conscious things can differ from those of unconscious things. For conscious things have this habit of sometimes going around worrying about questions of 'self', etc. – not all the time, no doubt, but at least at critical moments when the matter is brought up with them. This is, in a sense, often the way in which one is persuaded that some other object *is* actually conscious! Of course, it would be a simple matter to program a computer merely to go around muttering 'worry, worry, worry . . . self, self, self' . . . but it is hard to see why natural selection should trouble to produce creatures which do such things it if were not, in some sense, an important ingredient of consciousness itself – a consciousness that is no doubt naturally selected, for quite other reasons.

What is Consciousness Needed for?

Though it is hard to be precise about the distinction between that kind of activity which seems to require consciousness and that which does not, I should like to list some suggestive terms that we do, in the ordinary way, use for one such kind of activity or the other:

consciousness not needed	*consciousness needed*
'automatic'	'common sense'
'following rules mindlessly'	'judgement of truth'
'programmed'	'understanding'
'algorithmic'	'artistic appraisal'

I suppose, to a neurophysiologist, one might try to suggest that the activities in the left-hand column might be carried out by the cerebellum (see Rodolfo Llinás' chapter), whereas those on the right require some activity of the cerebrum. I doubt that it is quite as simple as that, but there is likely *some* truth in it. (Perhaps one could describe what an 'algorithm' is to a neurophysiologist by saying that it is an activity that could be carried out by a cerebellum!)

The distinction is not, however, even as clear-cut as I have presented it. It seems to me to be manifest that there are many unconscious factors which enter into our conscious judgements. Matters of experience, prejudice, intuition, and even our normal use of logic, seem to be involved in our judgements in ways that we are commonly not aware of. Perhaps it is the case, as many would suggest, that even the conscious parts of these judgements are still algorithmic, but just incomparably more complicated – or 'higher-level' – than are the unconscious ones.

Personally, I would strongly doubt that consciousness can be just a reflection of the complexity of an algorithm. I am also doubtful that the term 'higher-level' is, in itself, greatly helpful as a means of distinguishing the putative 'conscious' algorithms from the unconscious ones. For the 'level' that one assigns to a logical operation is largely a matter of one's interpretation of it – or of the 'meaning' that one attributes to it (that is, semantics) – rather than something that the computation itself intrinsically possesses.

With regard to the question of complexity, there is a property of

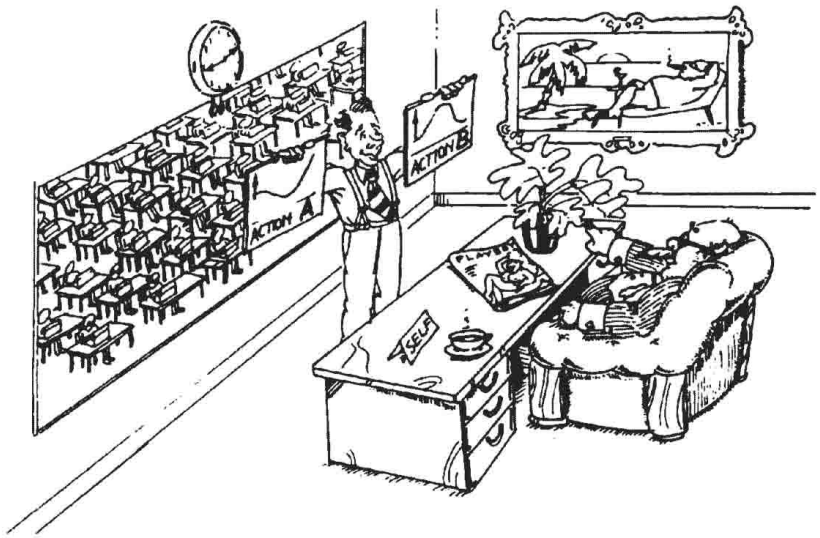


Figure 18.3 The conscious self *seems, in some respects, to be like the chairman of some large corporation who is presented only with highly processed and simplified data.*

conscious thinking that I have always found puzzling, namely that things *seem* to need to be very *simple* in order that our conscious judgements can take effect. Somehow, the complicated operations are all done for us unconsciously and only the final ruling between some simplified alternatives seems to be carried out by the conscious mind. It is rather like a chairman of a large corporation, who insists on being presented with highly processed and simplified data, and then makes all the 'important' decisions. I do not suppose that it is really quite like that. There is somehow a lot of information in those seemingly vague and 'over-simplified' sense impressions ('*gestalts*') that one's consciousness is continually being presented with (or perhaps actually 'is'). Yet there *is* something about having to keep thoughts simple in our conscious thinking. I am continually struck by such commonplace statements as: 'how can I be expected to think of more than one thing at a time?' Somehow it is intrinsically difficult – and even regarded as obviously so – to keep more than one thing consciously in mind at a time. This is a very puzzling feature of conscious thought, if consciousness is to be merely a matter of complication, since there is no difficulty at all about keeping a huge number of vastly complicated *unconscious* processes going on at once, as is continually the case in one's ordinary activity. But *consciously*, one seems to be able to keep only one thing in mind at a time. It would appear that to deal with several things at once the conscious mind needs to flip backwards and forwards between them rather than deal with them concurrently. There is a curious 'oneness' or perhaps 'globality' of conscious thought which appears to be among its characteristic features. I shall have a comment to make later of possible relevance to this.

Computer Models of the Brain

To see what all this might have to do with the suggested computer models of the brain, I shall need to examine briefly the physiological basis of such

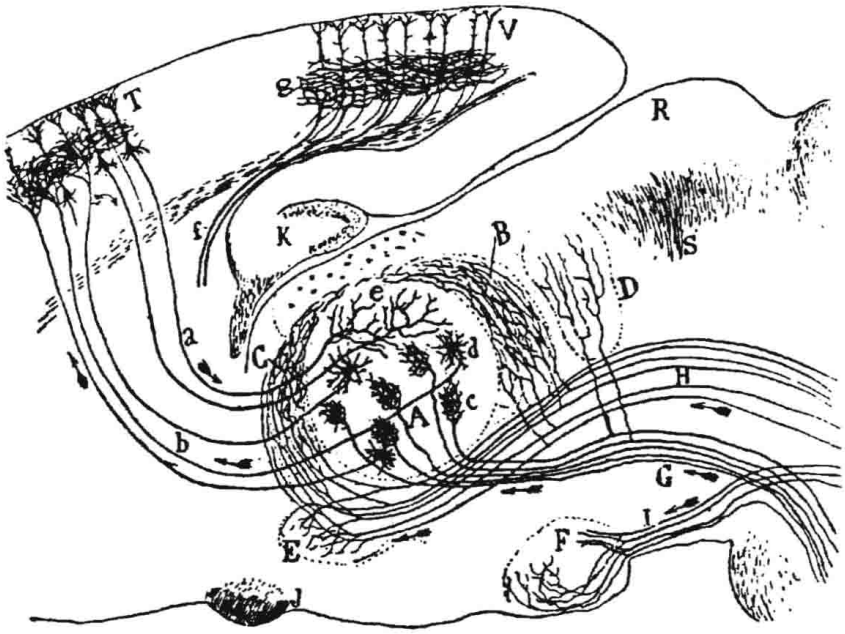


Figure 18.4 Huge numbers of nerve fibres connecting a great many neurons. This diagram (1911) by the Spanish anatomist Santiago Ramon y Cajal shows one side of the brain, cut through from back to front. It illustrates a tiny proportion of the different types of nerve cells, with fibres carrying in sensory information and carrying out motor signals.

models. The picture that is normally presented involves huge numbers of nerve fibres connecting a great many neurons; here the nerve fibres are modelled as wires carrying electric currents and the neurons as the logical switching units that occur, in computers, at the junctions of such wires – and which provide the logical (and hence also arithmetical, etc.) operation of the computer. The present view seems to be that single neurons do not correspond just to the most basic logical units ('and', 'or', 'not', etc.) but to some more complicated combinations. This would not affect the general picture. However, there does seem to be an additional ingredient in that the firing of neurons appears to involve a *random* element. What comes out is not just a function of what goes in, but all one can say is that there is a *probability* so-and-so of such-and-such an outcome, for any given input signal. Thus one seems to have, in effect, a Turing machine with a *randomizer* instead of a Turing machine proper.

However, there are mathematical results which say that the randomizer does not really give us anything new in the sense of being able to compute things that one could not compute without the randomizer. In any case, what is in practice normally done in modern computers, when some random ingredient is needed in a calculation, is to supply what are called 'pseudo-random' numbers – the results of some calculation which are, to all intents and purposes, indistinguishable from 'truly' random numbers. Alternatively, one could incorporate some physical random input such as the decay times of atoms in some radioactive source. It seems hard to believe that this would make any difference in practice.

All this seems to suggest that the modelling of the brain by a Turing machine is a reasonable procedure. Indeed, I think that it *is* a very worthwhile thing to try to do. There are, however, certain objections that

one can make to taking this analogy too literally. In the first place, a Turing machine needs an infinite store, whereas the capacity of the brain, though very large, is clearly finite. On the other hand, one could also allow an unlimited supply of pieces of paper for writing notes on. Of course one could query whether in practice this *is* strictly unlimited (which it is not). But the real limitation is likely to be the *time* it takes to do a computation, not the number of pieces of paper that are allowed to be used. A computation that takes $10^{10^{10}}$ years to perform using some well-defined algorithm is not going to be much use in practice – and algorithms can, indeed, often be absurdly slow. The general theoretical discussion of computability by algorithms is very much concerned with what can be done in principle and not really with how long (in the sense of how many steps) it takes to carry out a computation.

Complexity

This leads us into another area of mathematics known as *complexity theory*, which *is* concerned with this very type of question. Given a well-defined class of mathematical problems for which there does exist an algorithm for solving them, and for which the 'size' of the problem is characterized by some positive n, we may ask how long the algorithm takes as a function of n. For example, the time might be roughly proportional to $\log n$, or n, or n^3, or 2^n, or $n!$, or n^{n^n}. In the last few of these cases the algorithm would rapidly become useless for moderately large values of n. In complexity theory, one of the aims would be to find algorithms in which this function of n does not grow too rapidly, where some algorithm may be already known which is perhaps inefficient in this respect – or else to find absolute limitations to possible improvements of this kind. Clearly these questions have some importance in the matter of what can be computed *in practice* by either a human or electronic computer. But even complexity theory is not exactly concerned with the practicality of algorithms, since it deals only with their efficiency in the limit of very large values of n.

The excessive slowness or memory requirements of certain algorithms would seem to be to the disadvantage of the human who chooses to use them, more than they are to the electronic computer, since inefficient 'brute force' operations are carried out much more efficiently and amenably by the latter. It is a more human quality to try to see one's way around having to do such things – for example, by finding shortcuts, or even by building computers to carry out such burdensome tasks!

Gödel's Theorem

Is there evidence, then, for ways in which the human procedure of conscious thinking might actually be *more* effective than the purely algorithmic computer picture of things? Is there even any clear evidence that conscious thinking is actually any *different* at all from something entirely algorithmic? My personal view is that there *is* such evidence, though none of it is uncontroversial. It would take too long to expound on these matters to any adequate degree, but I should like to make a few

comments about Gödel's theorem, which I believe to have some genuine relevance to this.

Gödel's argument is concerned with the question of the provability of mathematical statements by means of some pre-assigned system of precise formal procedures.[6] For definiteness, let us take these mathematical statements to be propositions about arithmetic – statements of the form: for all integers $a, \ldots, c$, there exist integers $d, \ldots, f$, such that for all . . . etc., then some arithmetical relation between $a, \ldots, c, d, \ldots, f, \ldots$ holds. The statements of 'Fermat's last theorem' or the 'Goldbach conjecture' have this general form. The pre-assigned proof procedures are supposed to have the property that their correct application or otherwise can be checked by some algorithmic procedure, that is, by a Turing machine. Assuming that these procedures are actually all valid (that is, that they lead only to *true* arithmetical propositions), Gödel's argument contrives to produce another arithmetical proposition that is *not* obtainable by the procedures, yet is manifestly *also true*!

What Gödel's argument *seems* to show is that the way in which we decide that mathematical statements are true is not algorithmic. It most definitely does show that there is no algorithmic procedure (a Turing machine) which can generate *all* the true propositions of arithmetic and no false ones. It is often argued, on the other hand, that there is no evidence that human mathematicians can do any better. Yet it is a remarkable fact that the very argument by which Gödel establishes his result provides a specific arithmetical proposition that any mathematician following the Gödel argument *sees* to be true. By following the argument, a *new* method of proof is seen to be valid, which goes beyond whatever collection of methods had been accepted previously.

But why, then, can one not simply get a computer also to follow this Gödel argument and itself 'see' the truth of any new Gödel proposition? The catch lies in 'seeing' that the Gödel argument, in any specific realization, has actually been correctly applied. The trouble is that the computer does not have a way of judging truth; it is only following rules. It does not 'see' the validity of the Gödel argument. It does not 'see' *anything* unless it is conscious! It seems to me that in order to appreciate the validity of the Gödel procedure – or, indeed, to *see* the validity of *any* mathematical procedure – one must be conscious. One can follow rules without being conscious, but how does one *know* that those rules are legitimate rules to follow without being, at least at some stage, conscious of their meaning?

This matter has a bearing on the 'simplicity' inherent in conscious thinking that I referred to earlier. I am not sure that 'simplicity' is really the correct term, but I am unable to think of an appropriate alternative. The 'simple' thoughts that I am referring to are often hard to formulate in a precise way. They are things that one would often be hard-put to program on a computer. 'Common sense' is a good example. The central idea underlying the Gödel argument is another. There are many things in mathematics and science that could come under this heading. It is a question of conceptual simplicity, perhaps, though specific realizations of the 'simple' ideas may well lead to extreme complication in detail (as they do with the Gödel argument).

Physical Laws

The main argument *against* human mathematicians being able to achieve what a computer cannot rests, as far as I can see, on the idea that it should be possible *in principle* to follow whatever physical processes are going on in a person's head by some precise computation, even if one does not know the details of which computation that might be. The argument goes beyond any specific model of nerves and neurons and refers directly to the physical laws that are operative on a human scale. One has the picture that the material of a brain is composed, after all, of a vast number of particles which move around and bounce off one another, all according to very precise laws. Some of the particles are charged and are deflected in precise ways by a pervading electromagnetic field; and this field itself propagates according to the marvellously precise Maxwell equations. Is this not a highly mechanistic picture in which the future state can be computed – albeit only in principle – from the state at any earlier time? The time-evolution of any system, whether it be a precisely constructed machine or a biological system, would appear to be adequately governed, at the relevant level, by the deterministic laws of classical physics. Are we not being presented here simply with a vast computing machine?

It seems to be the prevailing view that classical physics should be adequate for describing behaviour on the scale relevant to a human brain. Nevertheless, it could be the case that *quantum mechanics* is actually needed. In my own view, it is highly likely that quantum phenomena *are* actually involved in the workings of the brain. It seems to be the case that single quantum events can be responsible for the triggering of nerve impulses. (In the visual system, a sensation of light can be triggered by something like five photons being absorbed by the rod photoreceptors of the retina; but even just a single photon can produce a detectable electrical response in a rod and could suffice to produce a visual sensation were it not for an inbuilt mechanism for suppressing or ignoring signals of that sensitivity – otherwise one would be flooded with irrelevant 'noise'.) How might this 'computer picture' be affected if a quantum description needs to be adopted?

Standard quantum mechanics,[7] in the way that one has actually to apply the theory, employs two distinct modes in which a system can evolve. The first is Schrödinger evolution (or, using a different but equivalent form of description, Heisenberg evolution – the more general comprehensive term would be 'unitary' evolution). According to Schrödinger evolution, the state of a system evolves according to a perfectly well-defined deterministic law. Everything is governed by a differential equation, the situation being very similar to that for the classical Maxwell equations, and again the 'computer picture' would seem appropriate. However, to get agreement with observation, one must bring in the second mode of evolution, which is referred to as 'state-vector reduction' (or wave-function collapse). The theory is not at all clear about when this second mode must be brought in. Phrases like 'when an observation is made on the system' are invoked in order to give us some indication of when this should be. But it is not very clear what kind of systems have 'observer status' and can so set off the 'state-vector

A Brief History of the Development of Quantum Theory

Max Planck 1900	black-body spectrum requires that light be emitted and absorbed in units of energy $E = h\nu$ where ν = frequency, h = (Planck's) constant
Albert Einstein 1905	postulates that electromagnetic fields consist of photons and predicts photo-electric effect
Niels Bohr 1913	('old' quantum mechanics) explains atomic spectra on basis of assumption that angular momentum of orbital electrons is quantized in units of $h/2\pi$
Prince Louis de Broglie 1923	postulates wave nature of particles
Wolfgang Pauli 1925	exclusion principle for electrons
Werner Heisenberg 1925, 1927	matrix mechanics, uncertainty principle
Erwin Schrödinger 1926	Schrödinger wave equation
Max Born 1926	probability interpretation of wave function
Paul Dirac 1927, 1928	general transformation theory, equation for electron, beginnings of quantum field theory

reduction' mode of evolution. It is this second mode that involves all the probabilities and the indeterminism that many people find so troublesome. The lack of determinism also makes it clear that state-vector reduction is not really consistent with Schrödinger evolution. One might like to imagine that when the observer is also treated as being part of a quantum system, evolving according to Schrödinger evolution, then the second mode could be dispensed with. But this appears not to be so, because Schrödinger evolution is deterministic whereas state-vector reduction is not.

I do not propose to enter into the endless discussions that these questions lead into. But if, for the moment, we simply accept that there are these two modes of evolution, then where does this lead us for our 'computer picture' of the brain? It seems that the interjection of this second mode at various times introduces an essentially random element, so we are apparently presented not with a Turing machine model, but with a Turing machine with a randomizer.

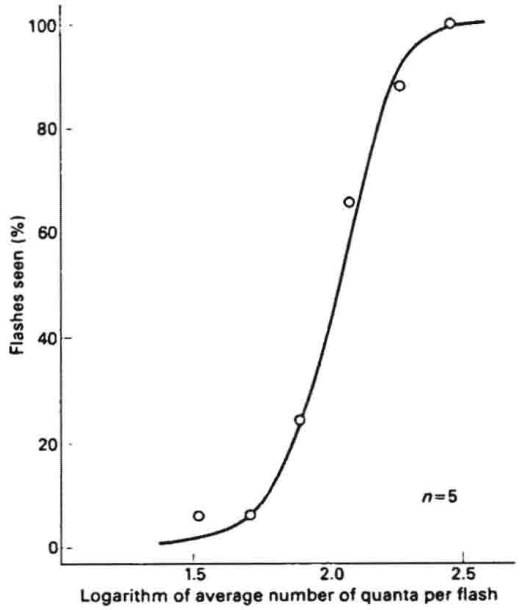


Figure 18.5 In 1942, the physicist Selig Hecht and his colleagues, S. Shlaer and M. H. Pirenne, reported the result of a remarkable experiment. They wanted to discover the smallest amount of light that can be seen by the human eye when fully dark-adapted. The observer (in this case Maurice Pirenne) looked at the dim flash and this graph plots the percentage of flashes that he saw according to the number of quanta or photons of light entering the cornea of his eye (2.0 on the horizontal scale indicates 100 quanta). A large fraction of the quanta are absorbed before they reach the rods and cones in the retina. The curve drawn through the points is derived theoretically on the assumption that only five quanta are needed for the flash to be seen. Since the flash covered a large number of rods in the retina, the chance of any one rod catching more than one of the five quanta was extremely low. So, the conclusion was that rods in the human eye must be capable of producing a signal on the capture of a single quantum of light (redrawn from the Journal of General Physiology, *1942, 25, 891–40).*

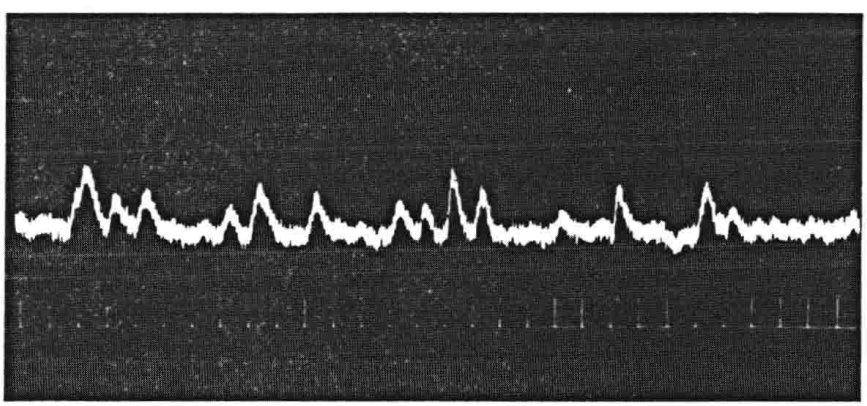


Figure 18.6 For these recordings of the electrical responses of a single rod from the eye of a toad, the light intensity was reduced to the level at which the rod absorbed only one photon, on average, *for each flash. Because of the probabilistic nature of light emission the actual number of photons hitting the rod would inevitably vary; sometimes none, often one, sometimes two, rarely three, and so on. The blips on the lower trace, occurring every 30 seconds, indicate the moment of each flash, while the upper trace is the amplified recording of electric current flowing through the rod. Sometimes there is no current, often a small response, sometimes one roughly twice as large, occasionally one about three times the size. Presumably these signals reflect the number of* individual *photons caught by the rod. Here is a direct demonstration of the ability of photoreceptors to respond to single photons of light (based on results published in the* Journal of Physiology, *1979, 288, 613–34).*

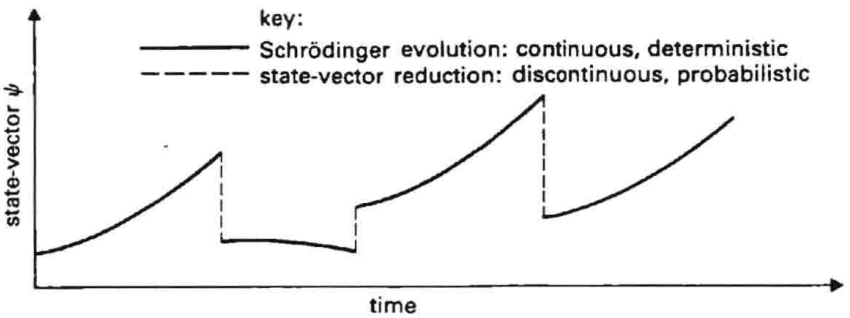


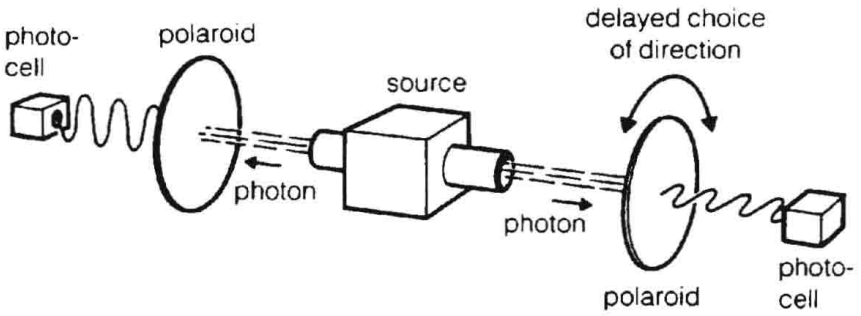
Figure 18.7 A strange picture of the evolution of the physical world – as described by conventional quantum mechanics.

Non-algorithmic Physics?

However, was I not arguing that the presence of a randomizer makes no essential difference to what can be achieved? I believe that this is so, and that the introduction of quantum mechanics, in the way that I have presented it would, indeed, make no essential difference. But in my view (and I am by no means alone in this) there is something very incomplete about quantum theory as it stands. It is a question of where the borderline falls between a quantum description and classical descriptions of things – a question of when a system should be described quantum-mechanically and when classically (the latter behaviour cannot properly be deduced from the former) – and of what to do with that borderline. I am of the opinion that a *new theory* is really needed in order to cope with such matters (and I believe that one can begin to make suggestions as to where to look for such a theory).[8] It seems to me to be quite on the cards that such a theory might involve operations with a non-algorithmic evolution!

Moreover, I think that this quantum–classical borderline could well be highly relevant in the workings of the brain. Whereas the firing of nerve signals can apparently be triggered by quantum events, it is the usual view that the signals themselves are 'large enough' to be treated as classical phenomena. However, the triggering of the next nerve signal could again involve single quantum events, and so on. I am not trying to say anything very specific here. But it seems to me to be highly likely that the quantum and quantum–classical aspects of physics ought not to be ignored in these discussions.

Quantum physics involves many highly intriguing and mysterious kinds of behaviour. Not the least of these are the quantum correlations (see box below) which can occur over widely separated distances.[9] It seems to me to be a definite possibility that such things could be playing a role in conscious thought-modes. Perhaps it is not too fanciful to suggest that quantum correlations could be playing an operative role over large regions of the brain. Might there be any relation between a 'state of awareness' and a highly coherent quantum state in the brain? Is the 'oneness' or 'globality' that seems to be a feature of consciousness connected with this? It is somewhat tempting to believe so. (Such a view might provide a way of resolving the 'teletransportation paradox', according to which a person is teletransported by photon signals to a distant



Quantum correlations: in the experiment of Aspect and his colleagues the polarizations of photons from a common source are measured some twelve metres apart. The photons behave as one still-connected entity up until they are detected, their polarizations being correlated with each other in a way which is inconsistent with their being separate objects (diagram highly schematic).

planet, but the instructions to destroy the 'original copy' of the person are disobeyed![10] Would such a hypothetical operation produce two absolutely identical versions, in different places, of a single *person*? What would happen to his consciousness? If a mental state is like a quantum state, then the paradox could be resolved. For *in principle* quantum states cannot be copied – or, rather, they can be copied *only* at the expense of destroying the original state.)

Finally, I should remark that even in classical physics, the relation between determinism and computability (algorithmic behaviour) is not at all as clear-cut as might be felt from the above discussion. Some mathematical results by Pour-El and Richards indicate that a non-computable evolution can sometimes arise even with very simple classical-type deterministic systems.[11] The precise status of these results in the present context is not clear to me. But they do at least highlight the important point that perfectly well-defined precise physical laws can sometimes result in non-algorithmic behaviour. Deterministic behaviour does *not* at all imply algorithmic (that is, computable) behaviour. (There are, indeed, many things in mathematics that are well defined but not algorithmic.) It is my guess that some non-algorithmic physical process may well be being harnessed in the workings of the conscious mind.

Notes and References

1 Turing, A. M. (1937) On computable numbers, with an application to the *Entscheidungsproblem*. *Proc. Lond. Math. Soc.* 2, 42.
2 Turing, A. M. (1950) Computing machinery and intelligence. *Mind*, 59; reprinted in A. R. Anderson (ed.), (1978) *Minds and Machines*. New York: McGraw-Hill.

3 Hofstadter, D. R. Jr. (1979) *Gödel, Escher, Bach: An Eternal Golden Braid* Hassocks, Sussex: Harvester.
4 For a collection of articles exhibiting a range of views on this question, see Hofstadter, D. R, Jr and Dennett, C. M. (eds) (1981), *The Mind's I.* For an extreme 'strong-AI' view, see Hofstadter's article: A conversation with Einstein's brain.
5 Searle, J., in ibid, also chapter 15, this volume.
6 See Nagel, E. and Newman, J. R. (1959) *Gödel's Proof.* London: Routledge & Kegan Paul.
7 See, for example, Polkinghorne, J. C. (1984) *The Quantum World.* London: Longman.
8 Penrose, R. (1985) Gravity and state-vector reduction. In R. Penrose and C.J. Isham (eds), *Quantum Concepts in Space and Time,* Oxford: Oxford Univ. Press. Press.
9 Aspect, A. and Grangier, in ibid.
10 See the introduction to ref. 4. and Derek Parfit's chapter in this book.
11 Pour-El, M. B. and Richards, I. *Adv. in Math.* 39 (1981): 215–39; 48 (1983): 44–74.

188
TILINGS AND QUASI-CRYSTALS: A NON-LOCAL GROWTH PROBLEM?

Introduction to the mathematics of quasicrystals. Ed. M. V. Jaric, 53–79. *Aperiodicity and Order,* Vol. 2, Academic Press, Boston, 1989

This is the only extended article that I have written on the subject of 'Penrose tiles' and quasicrystals, but see Chapters 64, 111, and 256.

812 Tilings and quasi-crystals: A non-local growth problem? *Introduction to the mathematics of quasicrystals*. Ed. M. V. Jaric, 53–79. *Aperiodicity and Order*, Vol. 2, 1989

Chapter 2

Tilings and Quasi-Crystals; a Non-Local Growth Problem?

R. PENROSE

Mathematical Institute
Oxford University
Oxford, England

Contents

1 Introduction 53
2 The Non-Locality of Mistakes 57
3 Fibonacci Sequences 71
4 Non-Locality of Fibonacci Sequences 76
References 78

1 Introduction

Shechtman and his coworkers announced their discovery of an icosahedral quasi-crystalline phase of aluminium-manganese alloy in 1984, but interest had been expressed a good deal earlier in the theoretical possibility of actual quasi-crystalline substances having icosahedral symmetry—with a structure perhaps similar to that of the tiling patterns that I had previously found (Penrose, 1974, 1979; *cf.* Mackay, 1982). Indeed, in the late 1970s and early 80's I had often been asked to give lectures on these tiling patterns, and a question frequently posed to me after the lecture might be: "Does this not mean that there is a whole new area of crystallography opening up, with pentagonal and icosahedral symmetry allowed?" My normal response

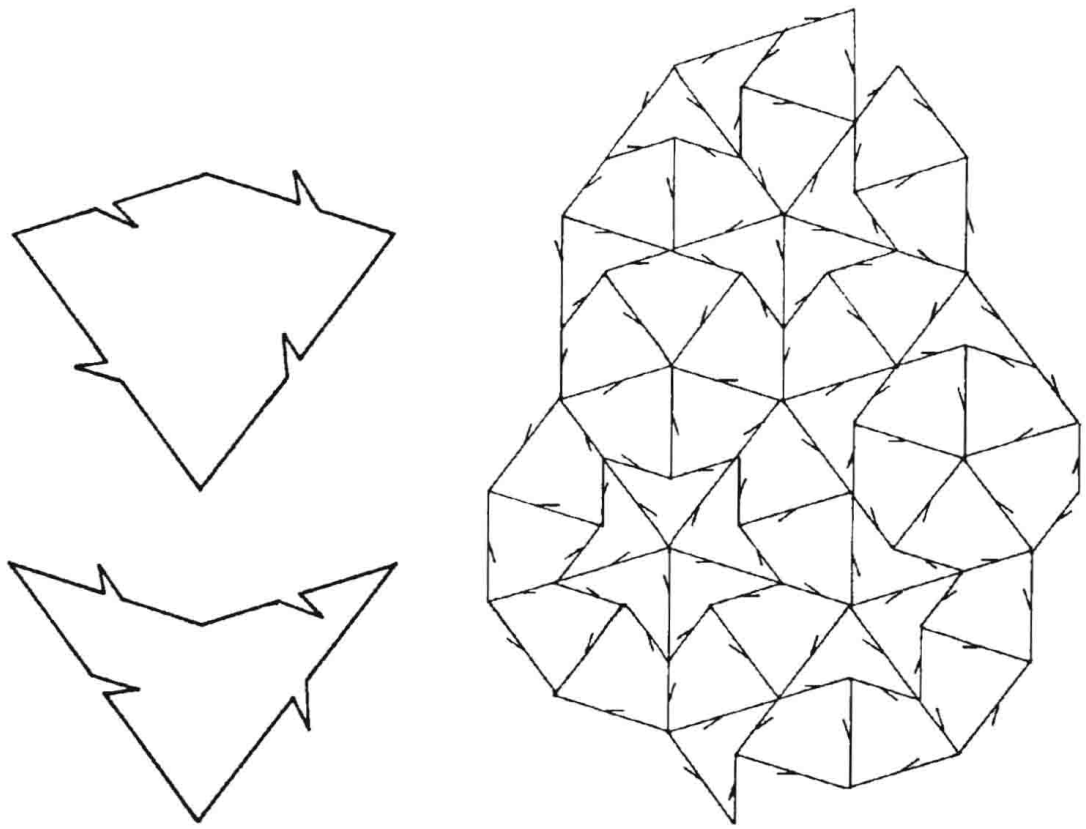


Figure 1. Left: A Kite (above) and a dart (below). Right: A portion of the plane tiled with kites and darts.

would be: "In principle, yes; but how on earth would Nature do it?" I had in mind that such substances might, with great difficulty, be perhaps synthesized artificially, but the problem confronting the *spontaneous* growth of large regions of such quasi-crystalline five-fold-symmetric substances had seemed to me to be virtually insurmountable.

To appreciate the problem that I had envisaged, imagine a plane tiling—say a kite and dart pattern (Penrose, 1979, 1987; *cf.* Gardner, 1977; Grünbaum and Shephard, 1987; see Fig. 1)—to be built up one tile at a time, as a jig-saw puzzle might be assembled. (An equivalent rhombus pattern would also do, but kites and darts are clearer, so most of my discussion will be formulated in terms of kites and darts. The modifications necessary for the rhombus case will be indicated where appropriate.) If the tiles are simply placed one after the other, wherever they will fit (the correct matching rules being assumed), then it will usually not be long before a configuration is reached in which there is some gap which cannot be filled in any way by either a kite or a dart. In order to continue such a pattern further, it would be necessary to remove a number of the tiles already placed, and then put

 Tilings and quasi-crystals: A non-local growth problem? *Introduction to the mathematics of quasicrystals*. Ed. M. V. Jaric, 53–79. *Aperiodicity and Order*, Vol. 2, 1989

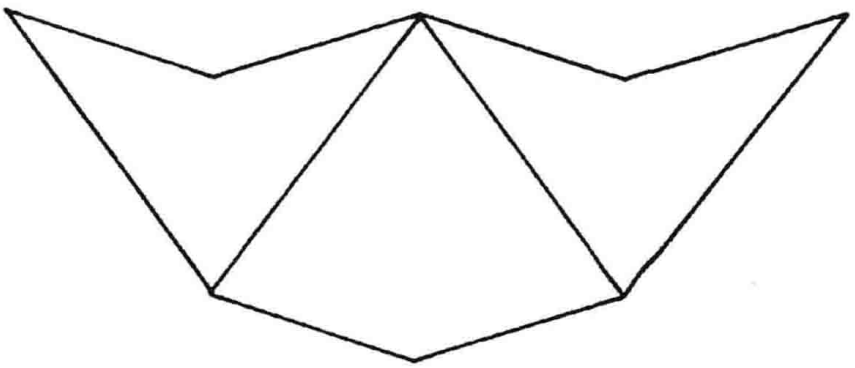


Figure 2. A simple "mistake".

them back in a different arrangement. Now, suppose that, in the assembly, the tiles are not simply placed so blindly—where originally the only criterion used had been whether or not they will fit locally—and, instead, some attention is now paid to the locations of tiles not immediately adjacent to the one being placed. For example, account may be taken of the locations of tiles *two* tiles away from the tile being placed. In this way we may avoid making "mistakes" of the kind illustrated in Fig. 2, where two darts and one kite are placed in a way that provides a "locally legal" pattern, but which cannot be continued to fill the whole plane (as can be readily seen). This enables larger patterns, on the average, to be constructed—until some arrangement is again almost certainly formed which cannot be continued further without gaps. If account is taken of the locations of tiles that are *three* tiles away from the one being placed, then even larger regions than before will be produced, on the average, before trouble tends to arise. But still non-continuable patterns will be almost certain to arise if the assembly merely proceeds in this largely random fashion. The conclusions would be similar if account is now taken of the locations of tiles that are *four* tiles away—or, as we shall see, with any other fixed bound on the distance away.

In my lectures on this topic, I usually illustrated this point with a slide showing the pattern of Fig. 3. I had made this slide from a picture that I had drawn out, without realizing, at the time, that the tiling was incorrectly assembled. The tiles are "bird" versions of the kites and darts, with markings on the back to indicate the relationship between these ("bird") kites and darts and a pentagonal type of quasi-periodic tiling that I had found earlier. The error is within the left-hand part of the figure. When subsequently examining the slide I realized that there were two stars (pentacles), formed from the markings, that should not be as close to one another as they are. This implies that there must be something wrong in the assembly and, indeed, it must be the case that if one attempts to continue the pattern somewhat further on the left, there must come a point where the

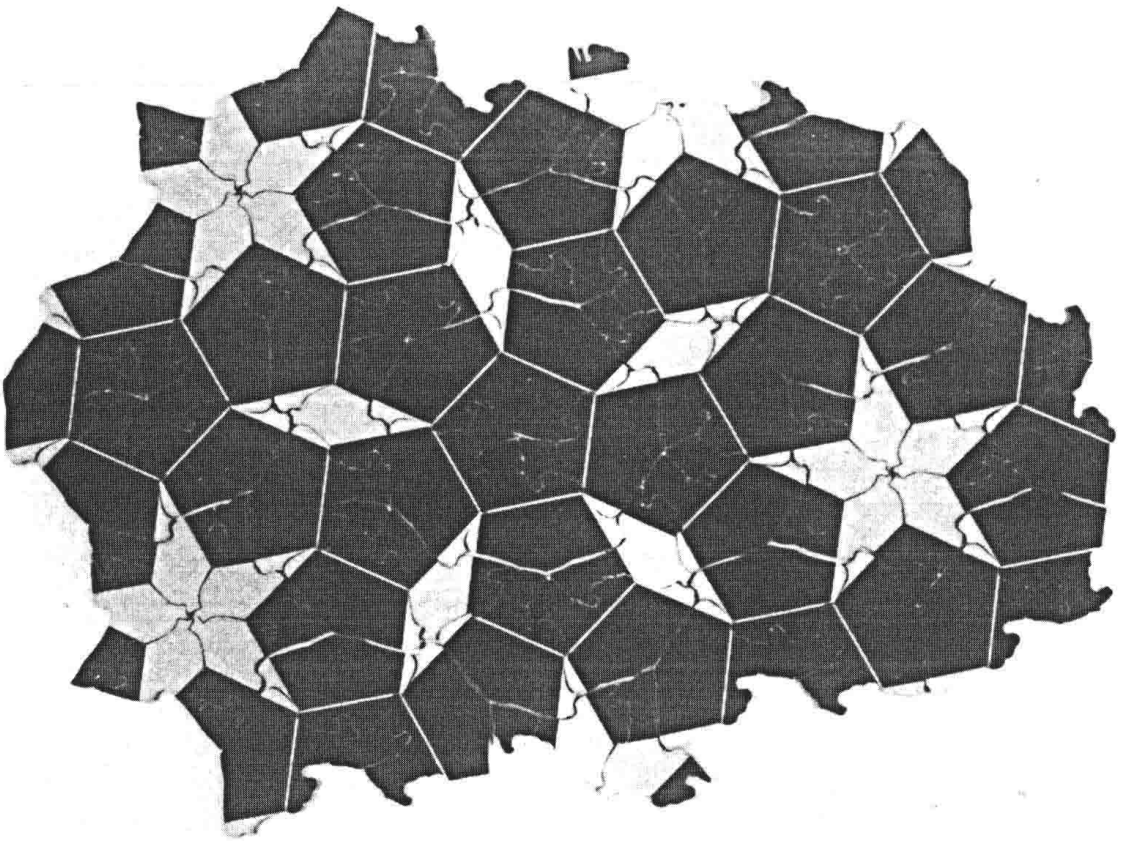


Figure 3. A tiling with (Escher-type) "bird" kites and darts (marked to bring out the relationship with an earlier pentagon tiling) which is "incorrect".

continuation becomes impossible, and one cannot proceed without leaving gaps. (The trouble is not immediately apparent; it becomes so only after various different attempts at continuation are made on the left.) In order to make a continuation which *will* eventually cover the whole plane, it is necessary first to reposition some of the tiles already placed—so that one or other of the stars on the left becomes only a three-pointed one (a "jester's cap"), as in Fig. 4.

How big a region must be examined at each stage in order that we may be sure that the pattern can be continued indefinitely without gaps? I have stated that there is *no limit* on the size of the regions which must be examined, at some stage or other. It is not hard to see why this must be so.

The kite and dart patterns are always based on a hierarchical arrangement. We may recall the *inflation* rules that Martin Gardner referred to in his 1972 Scientific American article (Gardner, 1977; *cf.* also Grünbaum and Shephard, 1987). In order to tile indefinitely large regions in a correct continuable fashion, one must, in effect, construct larger and larger "versions"

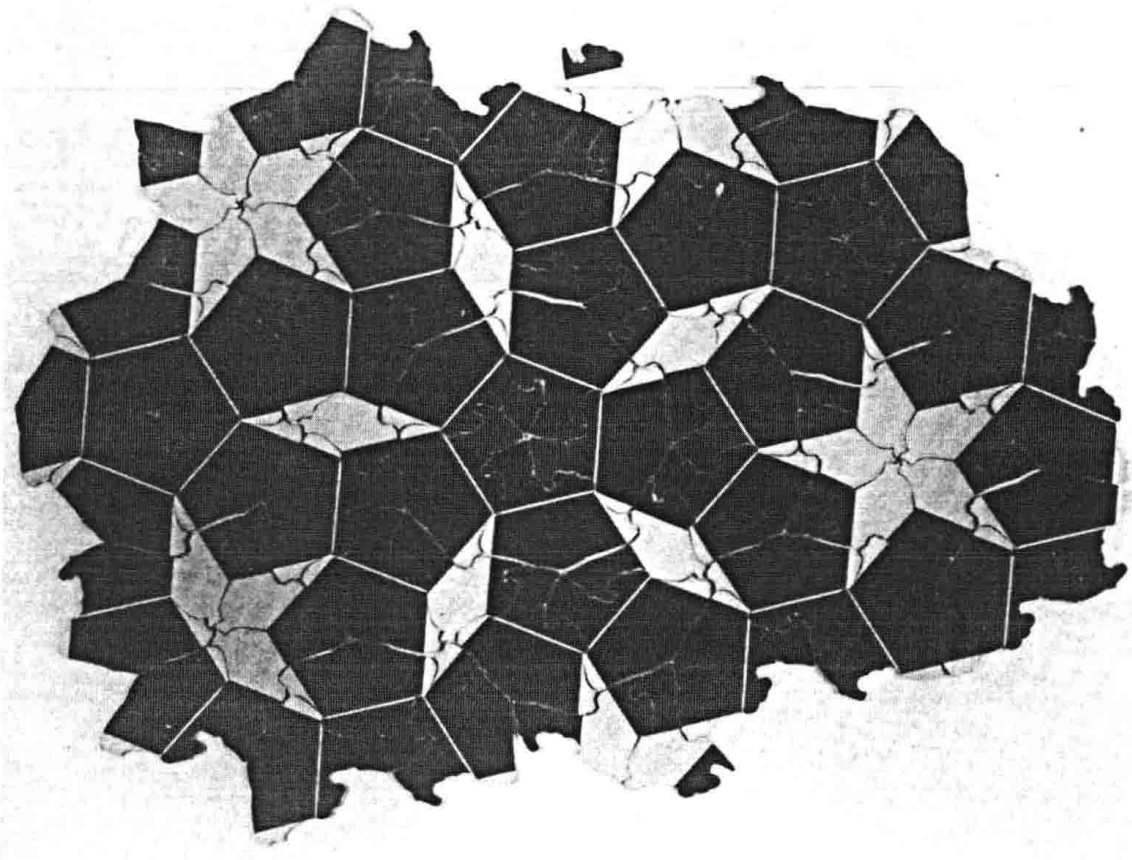


Figure 4. A corrected tiling with "bird" kites and darts.

of the original kite and dart tiles made up by the composition of greater and greater numbers of individual kites and darts (see Fig. 5). The corresponding "mistake" to the one just considered, for example, can be made again at every stage of the hierarchy (see Fig. 6). The larger the version of the "mistake", the harder it will be to see—in the sense of the farther away that one must look in the assembled pattern—that the pattern indeed cannot be continued without gaps.

In this article I shall try to provide some understanding of the non-local nature of such "mistakes." This might have significance in relation to the growth of quasicrystals—and to the question of how closely a physical quasicrystal might be able to approximate a three-dimensional version of these tilings.

2 The Non-Locality of Mistakes

For connected (and simply-connected) finite arrangements of tiles, I shall use the terms "legal" and "correct" in two clearly distinguished senses

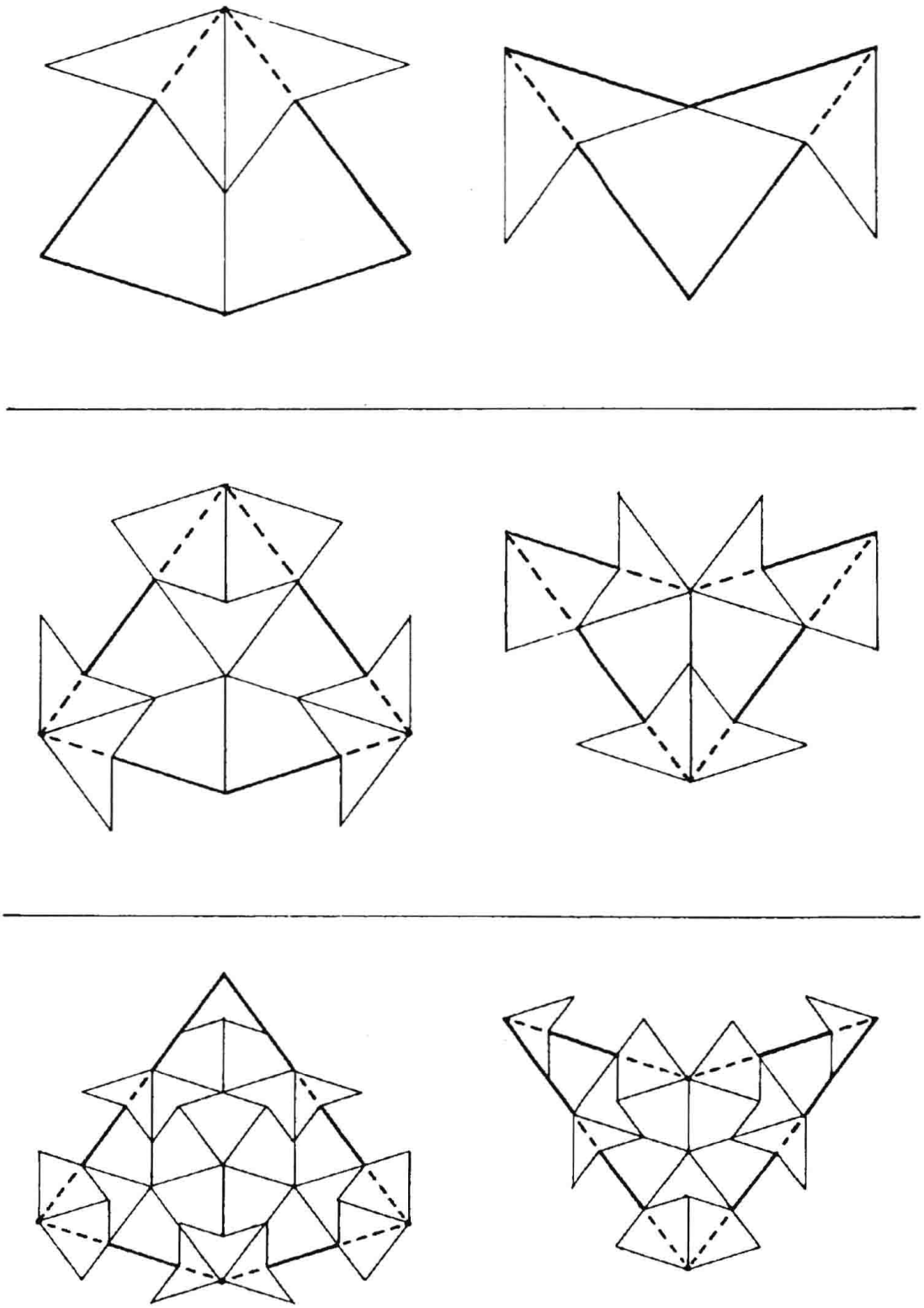


Figure 5. Three stages of hierarchy of kites and darts.

 Tilings and quasi-crystals: A non-local growth problem? *Introduction to the mathematics of quasicrystals*. Ed. M. V. Jaric, 53–79. *Aperiodicity and Order*, Vol. 2, 1989

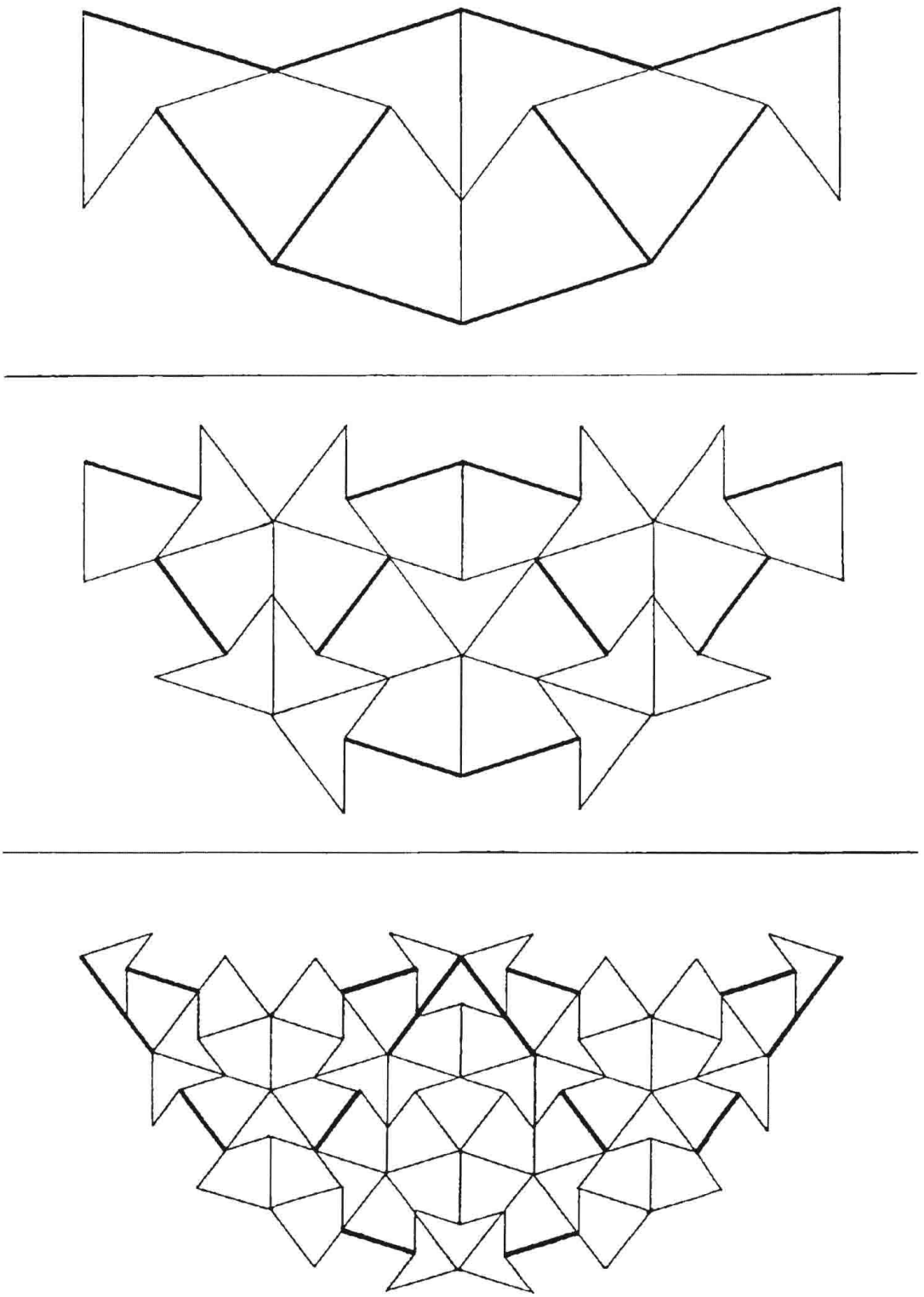


Figure 6. Incorrect tilings which are versions of the mistake of Fig. 2 at three different stages of hierarchy.

A *legal* arrangement is one where the tiles all have been fitted together locally without any gaps or overlaps, the local matching rules being all taken to be satisfied, but where a continuation to an infinite plane tiling is not necessarily possible. A *correct* tiling, on the other hand, is one which is a subset of a legal tiling of the entire plane. Thus, every correct tiling is legal, but some legal tilings are not correct, namely those which cannot ultimately be continued to cover the whole plane. I use the word "mistake," as above, for a placement of a tile which extends a correct tiling to a legal but incorrect tiling. Note that the tilings depicted in Figs. 2, 3, 6 are all legal but not correct.

Mistakes arise because there can be a choice as to whether it is a kite or a dart which must be placed along an unmatched edge at the outside of the growing pattern. Sometimes, because of the positioning of the neighbouring tiles, there is no such choice: just a kite or just a dart would fit as a legal placement. Sometimes, neither would be legal—but such a situation would arise only if a mistake had occurred at some stage earlier in the assembly, i.e., if the previous arrangement is not correct. Let us assume that *no* such mistakes have so far been made, so the previous arrangement is indeed correct. It then may be that *either* a dart *or* a kite can be correctly placed along the chosen unmatched edge, so that *both* placements can be continued to tilings of the entire plane. However, in other circumstances, it will be the case that only *one* of the two placements—either a kite or a dart—will be correct; but the other, though legal, is a mistake, and no continuation to the entire plane will turn out to be possible after that placement has been made.

In some cases, it may not be at all obvious whether or not such a placement is indeed a mistake. However, there is always an algorithm for ensuring correct placements: one can make use of the inflation rules referred to above. All patterns which continue to infinity must accord with these rules. We start with a correct configuration, which therefore has *already* been assembled in a way consistent with the inflation rules. Accordingly, the assembled kites and darts can be grouped together, by composition, to form larger kites and darts—where the darts must be bisected in order that the geometrical form of the kites and darts be preserved at each larger scale. The process is repeated with the larger kites and darts, so that these larger ones compose even larger kites and darts, and so on—until the original pattern, and its intended continuation, fall entirely within *one* large-scale kite or dart (see Fig. 7). There may be some choice as to how these groupings are made, but this does not matter. It merely reflects a choice which may be present in continuing the pattern.) Decomposing back down again we find a pattern of kites and darts on the smallest scale which includes the one we started with, and allows us to continue as we require, without mistakes. Repeating this procedure, we can continue indefinitely and cover as large a region of the plane as we please.

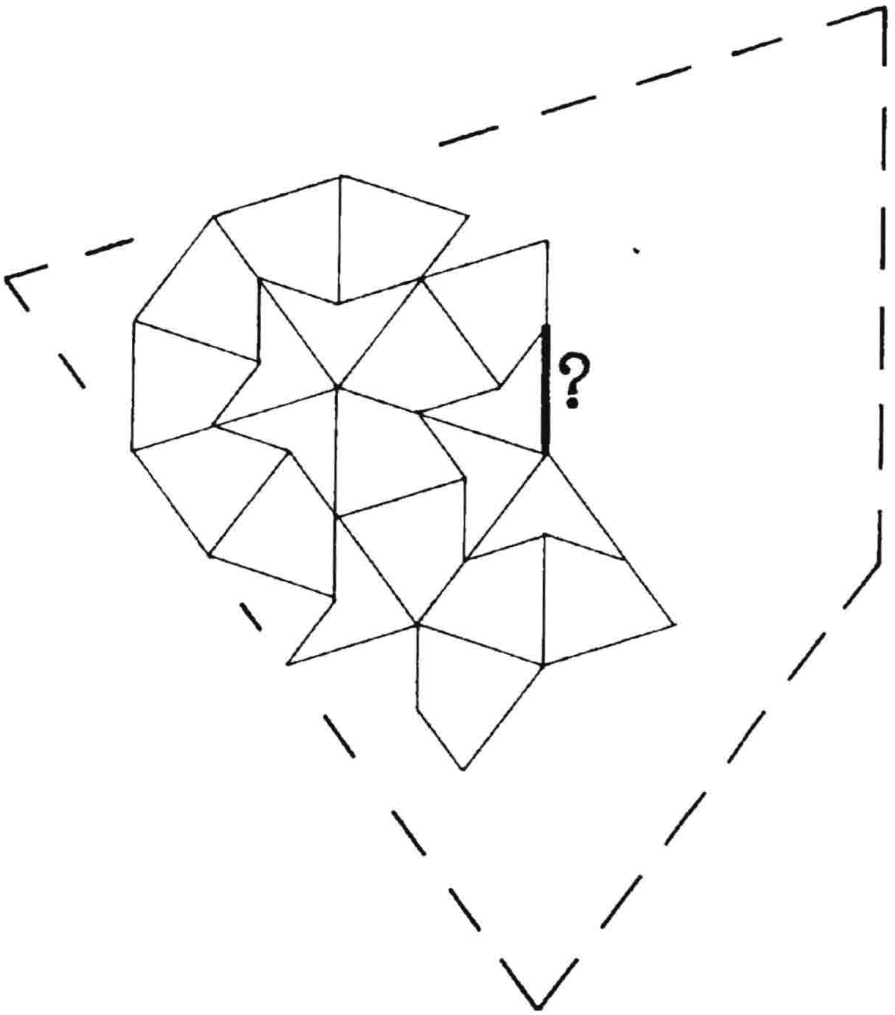


Figure 7. A method for deciding whether to place a kite or a dart at "?" is to proceed backwards along the hierarchy of Fig. 5 until all the tiles fall within one large-scale kite or dart, and then to proceed forward again by subdivision. In this case it turns out that only the dart is correct at "?".

Notice that this procedure is *not a local one*. We may need to examine the entire assembled pattern before we are assured how we may continue. As indicated above, there is *no* local algorithm, and this actually follows from the fact that, at every stage of the hierarchy, there is a version of the mistake of Fig. 2 that may be made (*cf.* Fig. 6). Let us explore this in a little more detail. First imagine that we have one kite and one dart joined together along one of their longer edges. If we wish to place another tile against the remaining longer edge of the kite, then we must make sure that this new tile is a kite and not a dart, otherwise we arrive at the forbidden configuration of Fig. 2. Now suppose that this original kite-dart pair is replaced by a much larger assembly of kites and darts, namely one obtained

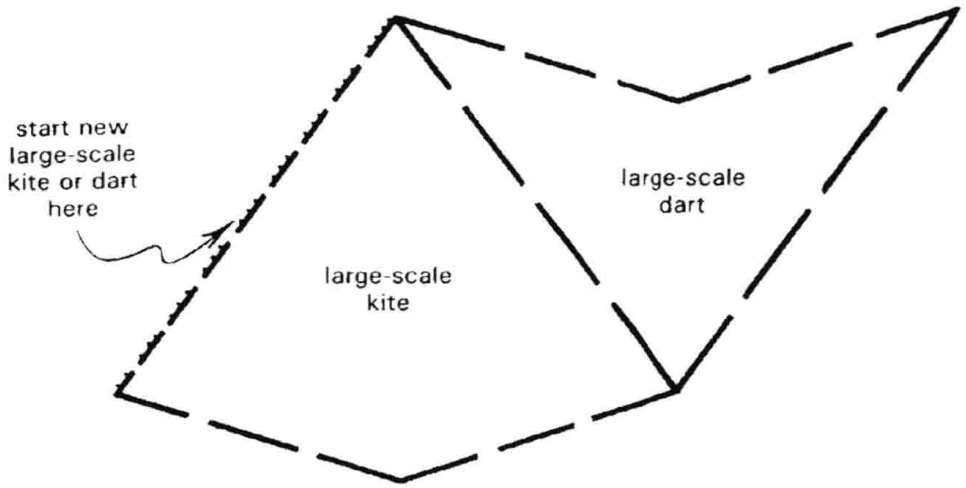


Figure 8. Suppose we wish to continue a correct tiling of a large-scale kite and adjoining large-scale dart by adding small kite and dart tiles along the sloping line at the left. If a new large-scale dart rather than a kite is assembled, this would be incorrect.

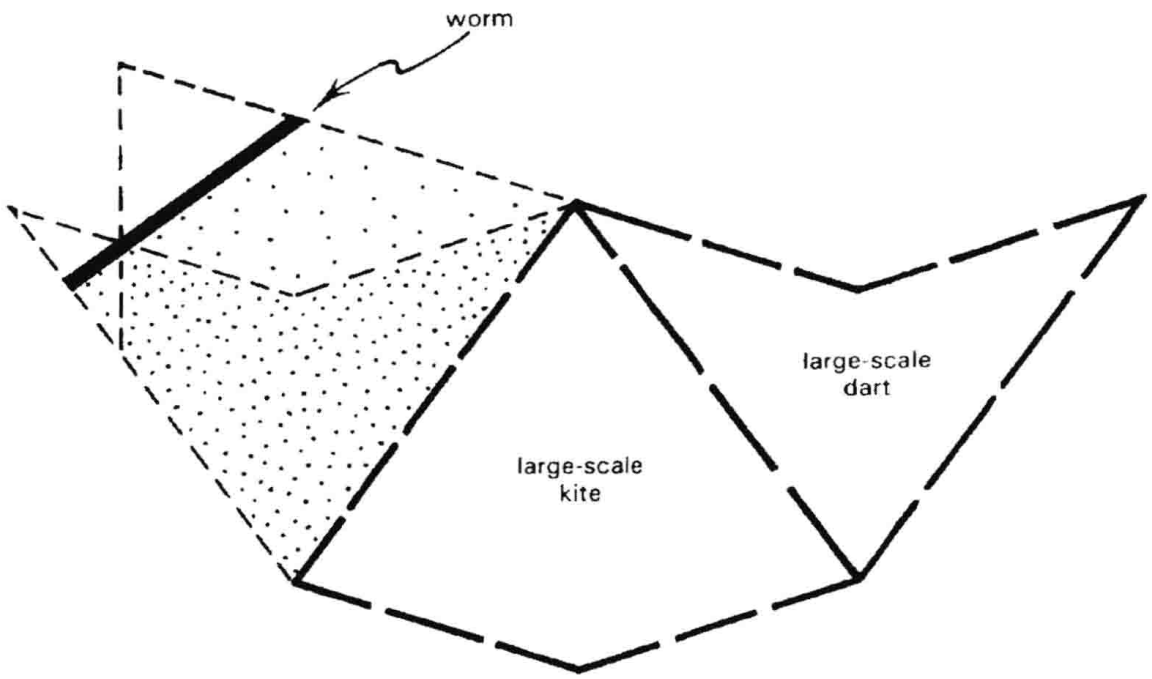


Figure 9. The difference between a new large-scale kite and a new large-scale dart, added on the left of the tiling of Fig. 8, does not arise until the *worm* (centered on the thick black line) is reached.

from the original pair by the above decomposition process repeated a great number of times. We now have a configuration which, on a large scale, resembles a kite and a dart joined along one of their longer edges, but in reality the whole pattern is made up of very many tiny kites and darts. Let us imagine that we are adding some additional tiny kites and darts in order to grow outwards from the unmatched longer edge of the large-scale kite (Fig. 8). We know that we must grow a new large-scale kite and not a large-scale dart. Otherwise we shall be producing a large-scale version of the mistake of Fig. 2 and 6. However, at first, the assembly of a large-scale kite turns out to be completely identical with that of a large-scale dart. If the individual tiny kites and darts are very small compared with the large-scale ones, then the large-scale kite that is being assembled will be practically complete before its detailed construction begins to distinguish it from a large-scale dart. There is a straight line, that I have indicated in Fig. 9, in the neighbourhood of which the differences between the large-scale kite and dart will first appear. The smaller that the tiny kites and darts are, the closer to this line the differences will lie. The line represents an example of what J. H. Conway has termed a "worm" (Fig. 10). The worm has to be turned over across its length in order to get from one large-scale pattern to the other. The way that the worms arise in the large-scale kite-dart pattern that I have just described is indicated in Fig. 11 and in Figs. 9 and 12.

We know that it is only the large-scale kite, in the position of Fig. 9, that allows an indefinite continuation to infinity without gaps; but this is the case only because of the positioning of the original large-scale dart. Had that large-scale dart been a large-scale kite instead, then *both* continuations would have been correct. But in order to *know* that this original large-scale dart was indeed a dart and not a kite, we should have had to examine the pattern as far as the corresponding worm at the *other* end of the pattern. Thus, the arrangement of tiny tiles in the worm at one end of the pattern will determine the correct arrangement in the worm at the other end. The entire arrangement in either worm is completely fixed by any *one* tile in the worm. Thus, the position of a single tile in the worm at one end, say at A in Fig. 13—assuming that it lies within a large-scale *dart*—will determine the positioning, say at B of the correct first tile to be placed in the worm at the opposite side of the pattern, namely that consistent with a large-scale *kite*. The alternative tile, placed at B, would be a mistake.

What will go wrong if we *do* allow ourselves to make this mistake? We can place tiles so as to fill up the entire interior of the triangle ABC, where C is the point at which the extended worms meet. However, at C itself it will not be possible to place either tile. The mistake at B is a mistake because of a tile placed at A, and the fact that it is a mistake becomes apparent only when the continuation has proceeded as far as the point C!

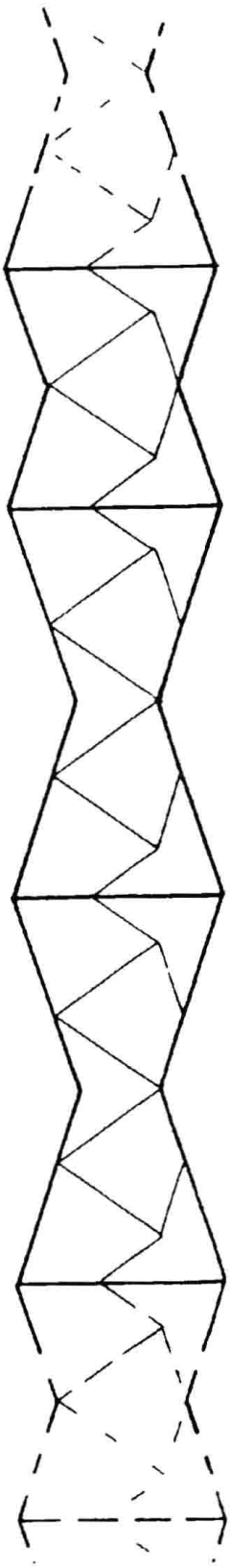


Figure 10. A worm (of kites and darts).

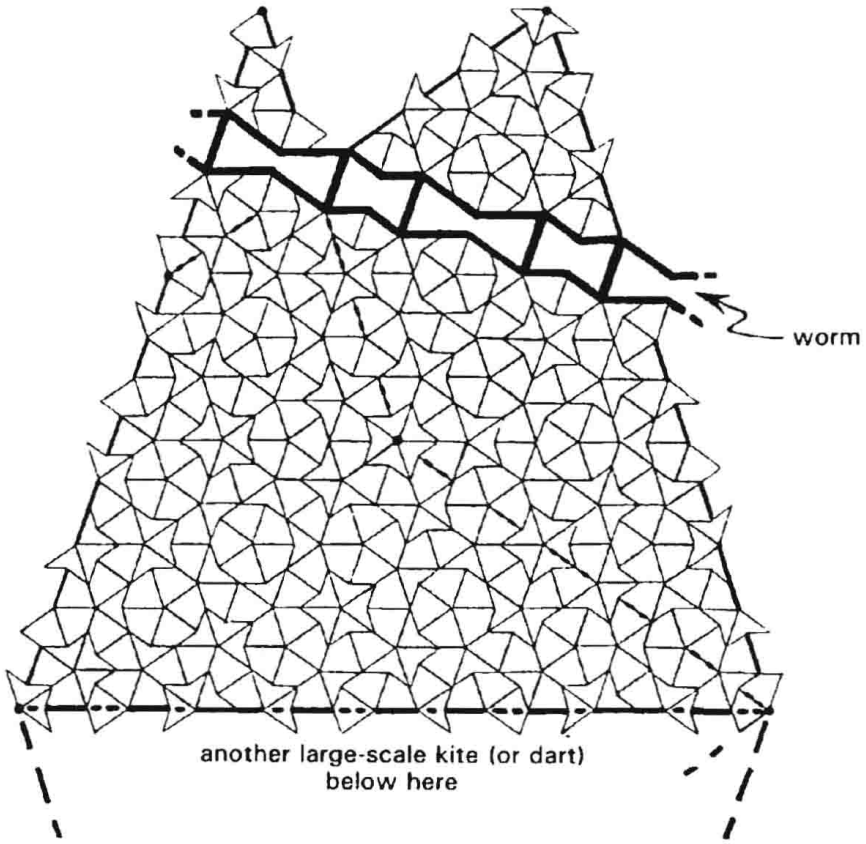


Figure 11. The detailed structure of a large-scale kite or dart, assembled on the long edge of another large-scale kite (or dart). Whether a kite or a dart is assembled, the detailed pattern is the same—except that the worm indicated has the opposite orientation in the two cases.

One can set up situations where the manifestation of a mistake occurs at an even more remote location than this. In Figs. 14a and 14b, I have shown how far one must proceed, when building out from a large-scale kite or a large-scale dart, before there is any choice in the correct placements of individual kites and darts. There is only one correct continuation of individual tiles out as far as the four boundary worms indicated. Each single boundary worm could be placed either way around; but for any pair of adjacent worms, one out of the three possibilities is illegal. For all *four* worms, for a large-scale kite it turns out that only *five* out of the sixteen possibilities will be legal, while for a large-scale dart, *seven* of the sixteen possibilities will be legal. The allowed possibilities are indicated in Figs. 15a and 15b, where the arrows show the direction in which each worm is to be oriented. (It should be noted that the relation between the worm orientation and the arrow reverses at each stage in the hierarchy.) To see how remote the manifestation of a mistake can sometimes turn out to be, consider Fig. 16.

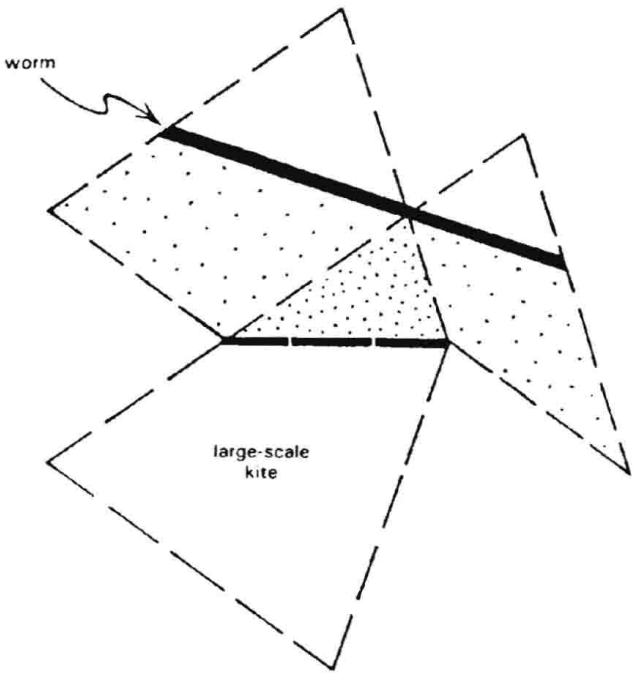


Figure 12. The detailed difference between a large-scale kite and a large-scale dart, when assembled on the *short* edge of another large-scale kite, is again along a worm.

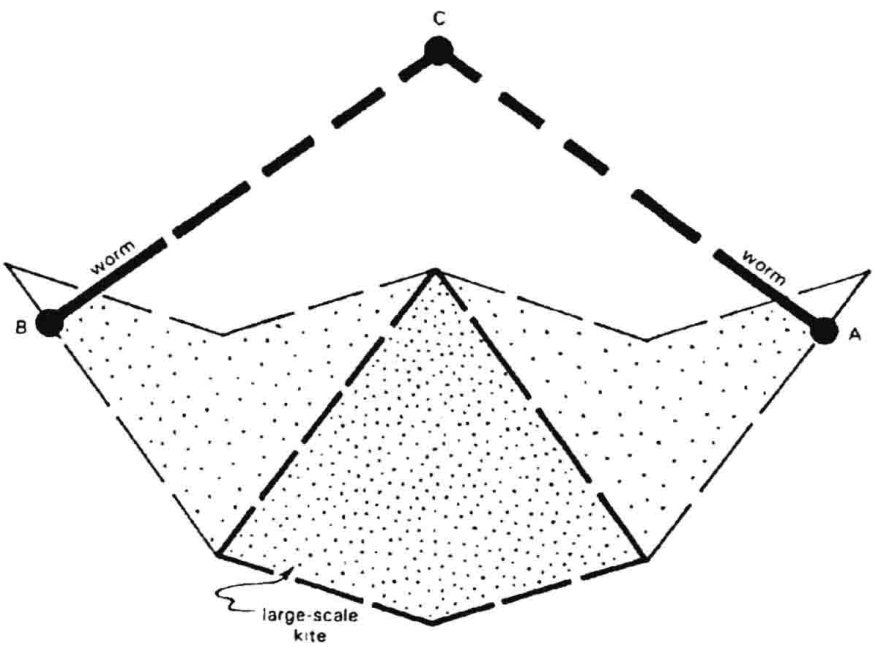
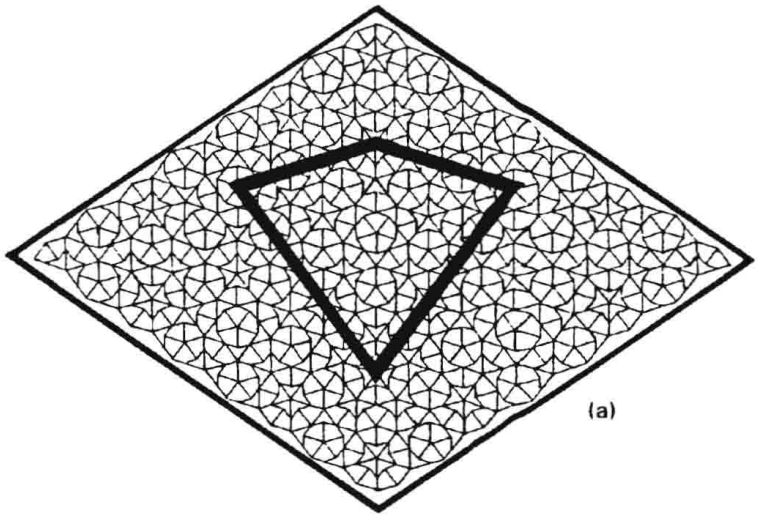


Figure 13. The placing of a tile at A which forces a large-scale dart on the right is incompatible with the placing of a corresponding tile at B forcing a large-scale kite on the left. The incompatibility of A with B does not emerge until C is reached—where the two worms meet.



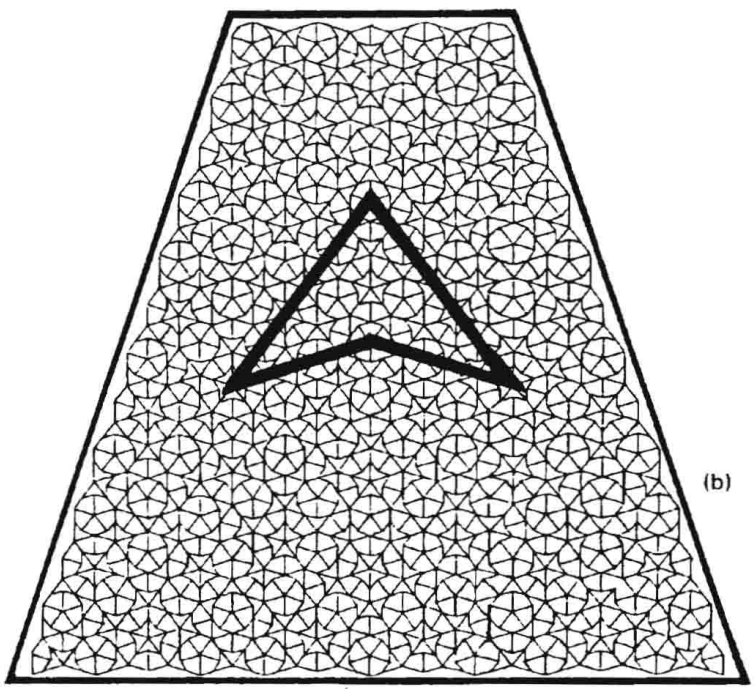


Figure 14. a. The central area is a large-scale kite, and a correct tiling proceeds unambiguously until the external quadrilateral is reached, whose sides are worms; b. the same for a large-scale dart.

Figure 15. a. Only five out of the sixteen possible worm orientations are legal for the quadrilateral surrounding a large-scale kite. (The relation between the arrow and worm-orientation reverses at each level of hierarchy.) b. The same for a large-scale dart: just seven out of the sixteen worm orientations are legal.

 Tilings and quasi-crystals: A non-local growth problem? *Introduction to the mathematics of quasicrystals.* Ed. M. V. Jaric, 53–79. *Aperiodicity and Order,* Vol. 2, 1989

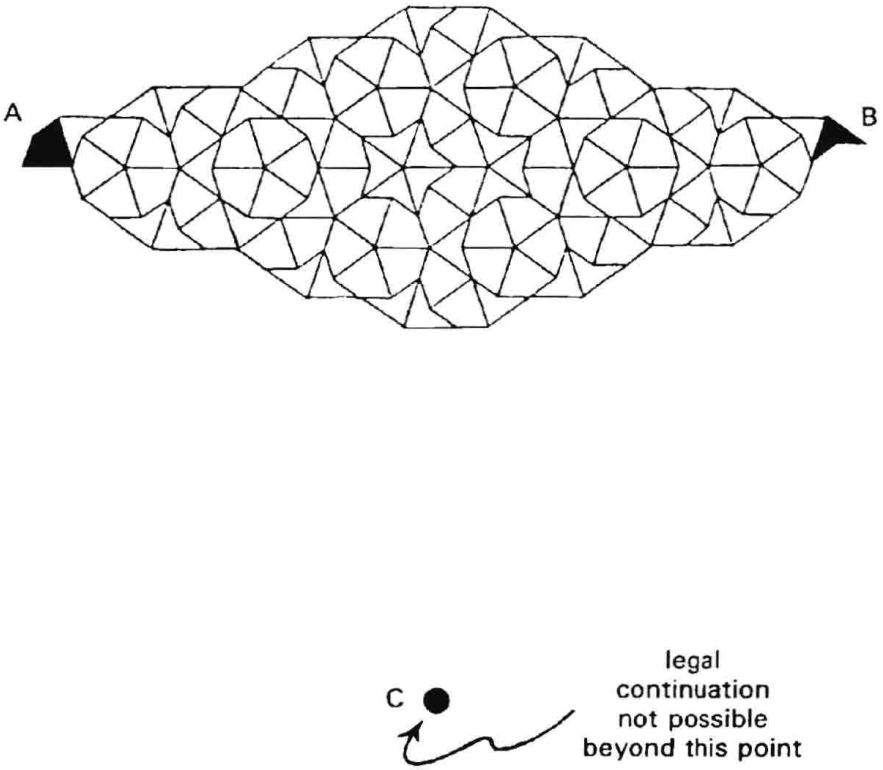


Figure 16. The placing of a kite at A is incompatible with a dart at B, whereas each would be correct without the other. The incompatibility only becomes evident when C is reached

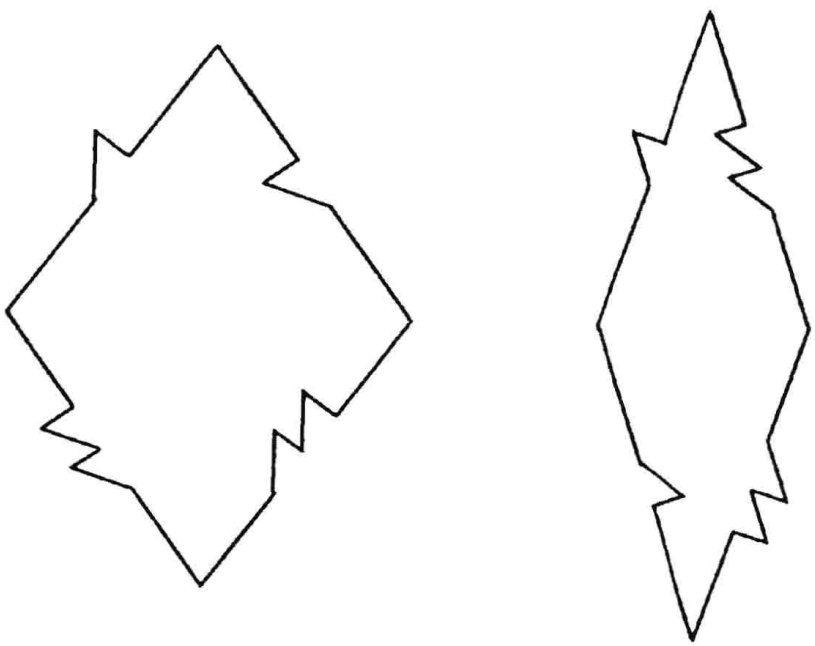


Figure 17. The rhombus tiles.

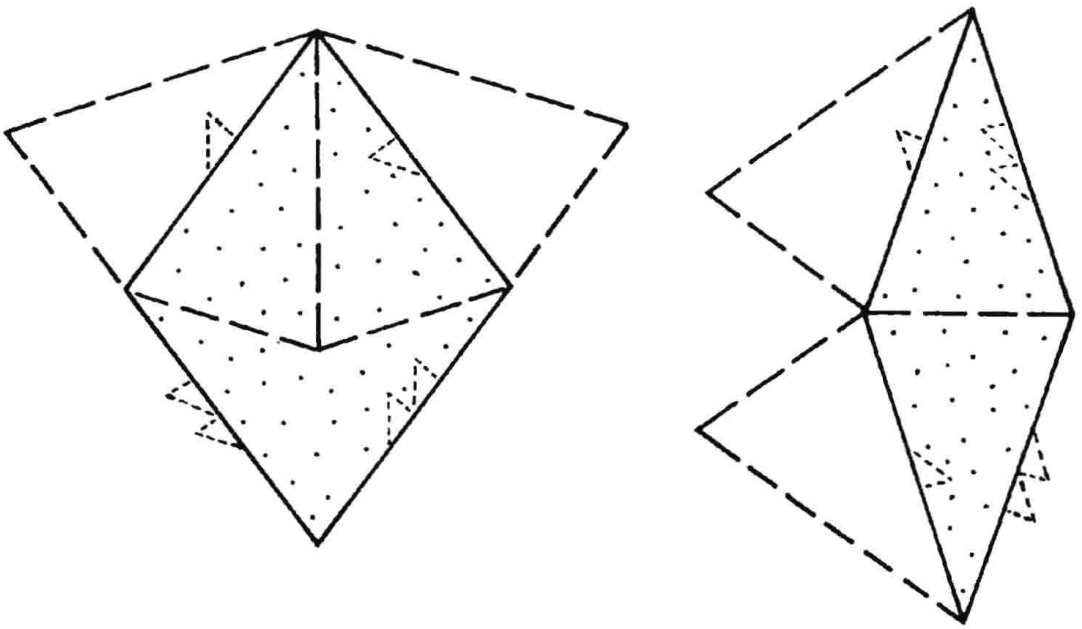


Figure 18. The relation between the rhombus pair and the kites and darts, at a particular stage of hierarchy.

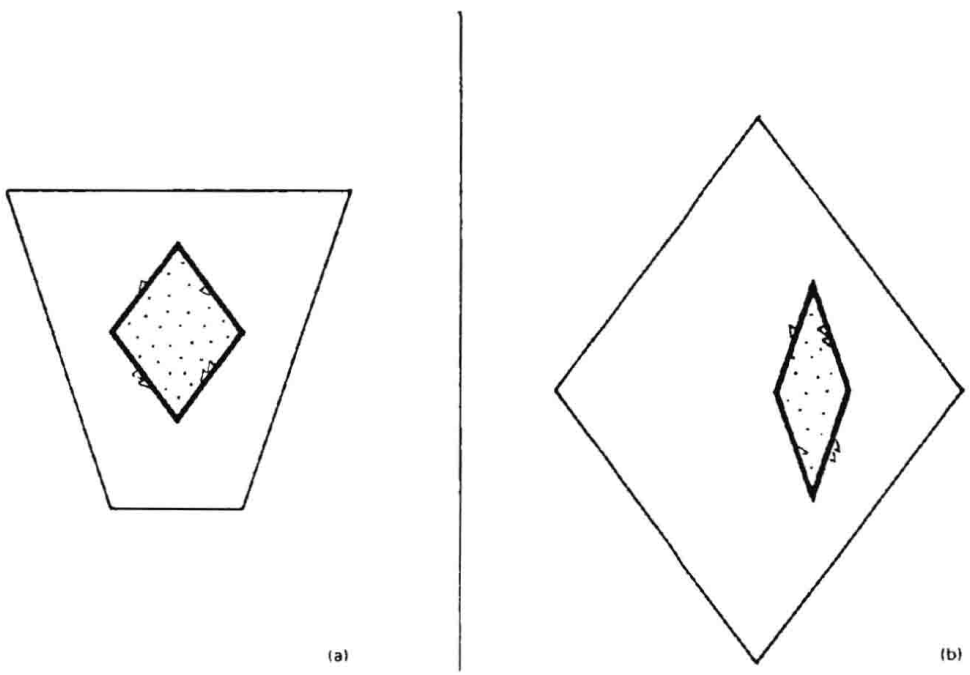


Figure 19. a. The central area is a large-scale fat rhombus, and a correct tiling proceeds unambiguously to the external region bounded by worms; b. the same for a large-scale thin rhombus.

The correct placement of a tile at B depends upon the tile which happens to be at A, and the effect of a mistaken replacement at B does not become manifest until the point C is reached!

Here I have been considering only kites and darts, but the situation with the rhombus tiles (with appropriate matching rules), as depicted in Fig. 17, is similar. The relation to the kite/dart tiles is shown in Fig. 18, at one stage of the kite/dart hierarchy. In Figs. 19a and 19b, I have shown the extent of the regions of forced correct tiling (as with Figs. 14a and 14b), now starting from each large-scale *rhombus* until the boundary worms are reached. The rules telling us which combinations of orientations of the worms are correct are the same as those for the kites and darts, as indicated in Figs. 15a and 15b (except that the relation between the direction of the arrow and the worm orientation is opposite, as between Fig. 15a and 15b.)

It may be remarked that it is always along worms that the differences between closely identical correct tiling patterns will occur. Striking patterns of straight lines can sometimes be seen when a transparency marked finely with either kites and darts or with rhombuses, is moved over another copy of itself until the patterns almost agree with each other. These straight lines are always worms.

3. Fibonacci Sequences

Notice that the worms are made up of two kinds of units, the long and the short. In the case of the kites and darts, these resemble long and short "bow ties" (*cf.* Grünbaum and Shephard, 1987), the long bow tie consisting of four kites and three darts, and the short one consisting of three kites and two darts. In the rhombus case, the worms are made up of two (non-regular) hexagons: a fat (i.e. "long") hexagon, built from two fat rhombuses and one thin one, and a thin (i.e. "short") hexagon, built from two thin rhombuses and one fat one (see Fig. 20). The long unit is, in each case, exactly the *golden ratio* ($= \tau = (1 + \sqrt{5})/2$) times as long as the width of the short unit. The sequence of long and short, for a correct tiling, always follows a strict rule, namely that given by a Fibonacci sequence (referred to as a "musical sequence" by Grünbaum and Shephard).

The rule for generating a Fibonacci sequence is a hierarchical one: start with 1 (representing a short unit) at the first generation; then replace 1 with τ (representing a long unit) at the second generation; then replace τ with 1τ at the third generation; then continue, making exactly these replacements again and again for each 1 or τ occurring in the entire sequence at each

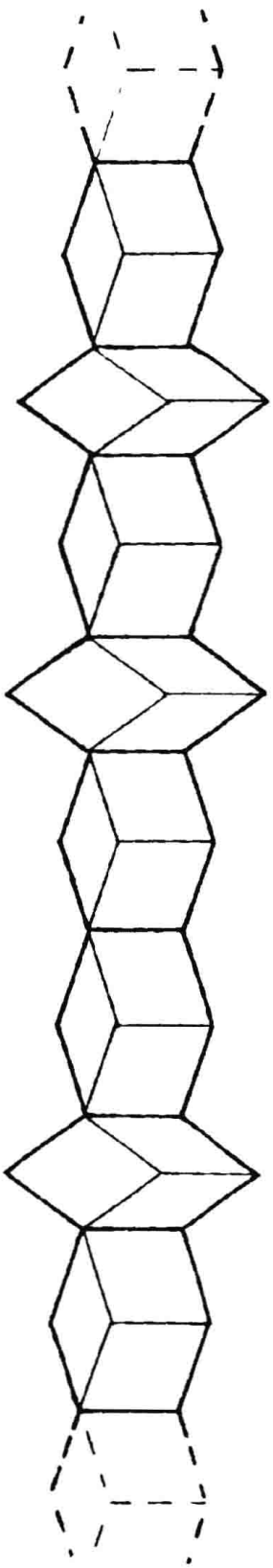


Figure 20. A worm of rhombuses.

generation. Thus we have:

$$\begin{aligned} &1 \to \tau \to 1\tau \to \tau 1\tau \to 1\tau\tau 1\tau \to \tau 1\tau 1\tau\tau 1\tau \\ &\to 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau \to \tau 1\tau 1\tau\tau 1\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau \to \ldots . \end{aligned} \tag{1}$$

The sequences that can arise in this way (perhaps infinite, in one or both directions), or any consecutive subsequence of such a sequence, will be referred to as a *Fibonacci sequence.*

To check whether or not a sequence of 1's and τ's constitutes a Fibonacci sequence, one can effectively apply the above procedure in reverse. One first ascertains that the original sequence never has two (or more) 1's in succession, and that the τ's occur either singly or in pairs; then one forms a new sequence in which each 1 is replaced by a τ and each pair of τ's by a 1, single τ's being ignored. Again, one must ascertain that the τ's occur either singly or in pairs in the new sequence (the 1's can now only occur singly), and the procedure to form a new sequence is repeated as before. If (and only if), in the continued application of this procedure, no sequence of more than two τ's ever arises (and no pair of 1's occurred initially) then the sequence is indeed a Fibonacci sequence.

As an example, let us apply this procedure to the sequence $\tau\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau$. We get

$$\tau\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau \to 1\tau 1\tau 1\tau\tau 1\tau \to \tau\tau\tau 1\tau, \tag{2}$$

which fails because of the triple τ, showing that the original sequence is not a Fibonacci sequence. But if we replace the first τ in our sequence by a 1, we get

$$1\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau \to \tau\tau 1\tau 1\tau\tau 1\tau \to 1\tau\tau 1\tau \to \tau 1\tau \to \tau, \tag{3}$$

showing that $1\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau$ is indeed a Fibonacci sequence.

Any worm which proceeds according to a Fibonacci sequence constitutes a correct tiling. (This is easily seen, because the inflation rule corresponds to the hierarchical rule for generating Fibonacci sequences.) However, a general sequence of 1's and τ's constitutes a tiling which is merely legal. Thus, for example, the worm corresponding to $\tau\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau$ would be legal but incorrect—so the tiles which build the first τ on to the rest of the sequence must involve a mistake. This mistake would become manifest only if the tiling is continued some way off to one side of the worm. We may note that the incorrect tilings of Fig. 6 arise, in effect, from the non-Fibonacci sequences 11 and $\tau\tau\tau$, respectively. In these cases the mistake would become manifest in the continued tiling some way above (although not below) the worm.

Figure 21. A test for a Fibonacci sequence (after de Bruijn). For each "1" we proceed vertically and for each "τ", horizontally. For a Fibonacci sequence, the points lie within a strip sloping at $1 : \tau$ and of such a width that it would just contain a lattice square.

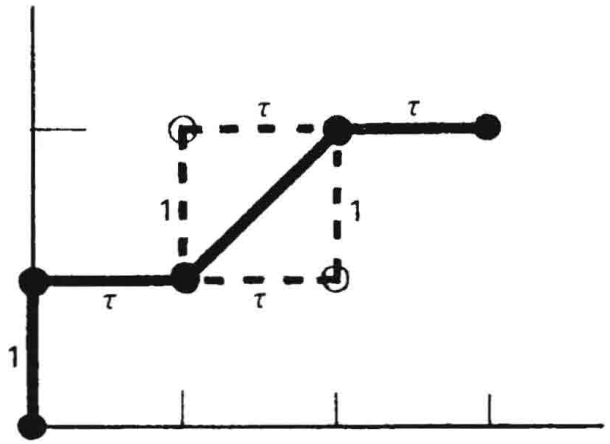


Figure 22. In case of doubt between 1τ and $\tau 1$ we proceed diagonally. This gives the quasicrystal "the benefit of the doubt".

Fibonacci sequences underlie correct kite/dart or rhombus tilings in a very basic way. (In particular, they govern the arrangement of wide and narrow Ammann bars; *cf.* Grünbaum and Shephard, 1987). If one examines such a tiling by viewing it obliquely from appropriate positions, with one's eye close to the plane of the tilings and off to one side, then one will observe such sequences of wide and narrow strips.

Electron micrograph pictures of icosahedral phase aluminium–manganese alloys will also exhibit such patterns of wide and narrow strips. It would seem reasonable to use the closeness with which these patterns accord with Fibonacci sequences as some measure of the closeness with which the atomic positions accord with an ideal icosahedral quasi-crystalline arrangement. A difficulty often arises in practice with this procedure, because it is not always clear whether a pair of adjacent terms in the sequence should be read as 1τ or $\tau 1$. To enable one to proceed nevertheless, one may, in such occasions of doubt, adopt a procedure which gives the quasi-crystal "the benefit of the doubt".

This procedure makes use of an alternative characterization of a Fibonacci sequence which I first learned about some years ago from P. Pleasants, who had adapted an idea of de Bruijn, 1981 to this situation.) Consider a sequence of 1's and τ's, say the one $1\tau 1\tau\tau 1\tau\tau 1\tau 1\tau\tau 1\tau$, considered above. We plot these as a zig-zag line on a square lattice, proceding horizontally by one step to the right for each τ, and vertically by one step upwards for each 1 (see Fig. 21). The condition for the sequence to be a Fibonacci sequence is that the resulting zig-zag line lie within a strip whose slope is $1/\tau$, and whose width is so that the opposite edges of the strip could just pass through the diagonal points of a lattice square. The way that one gives the quasi-crystal "the benefit of the doubt" is simply to plot the doubtful

cases, where one cannot decide whether a pair of adjacent terms is 1τ or $\tau 1$, as a diagonal to a lattice square, sloping upwards to the right (see Fig. 22).

My own conclusions based on a somewhat tentative analysis of various micrograph pictures of icosahedral phase aluminium-manganese alloys is that they tend to correspond remarkably closely to Fibonacci sequences, but the correspondence is by no means always perfect. Surprisingly, the type of "fault" that one tends to encounter is not a "glitch" here or there, but sometimes a slight, but significant, error in the slope of the strip. This is a genuinely *non-local* kind of error which I find to be perhaps as puzzling as an absolutely perfect "correct" assembly might be.

4 Non-Locality of Fibonacci Sequences

Suppose that we imagine a sequence of 1's and τ's growing in length, with one symbol being added at a time. Let us suppose that the growth proceeds randomly, but subject to some constraints. The constraints are to rule out incorrect sequences of symbols up to some prescribed length k—where "incorrect" means that it is incompatible with a Fibonacci sequence. The incorrect sequences are those containing the following:

$$
\begin{aligned}
&11,\\
&\tau\tau\tau,\\
&1\tau1\tau1,\\
&\tau\tau1\tau\tau1\tau\tau,\\
&1\tau1\tau\tau1\tau1\tau\tau1\tau1,\\
&\tau\tau1\tau\tau1\tau1\tau\tau1\tau\tau1\tau1\tau\tau1\tau\tau,\\
&1\tau1\tau\tau1\tau1\tau\tau1\tau\tau1\tau1\tau\tau1\tau\tau1\tau1\tau\tau1\tau1,\\
&\ldots\ldots;
\end{aligned}
\qquad (4)
$$

where each is, in effect, obtained from the previous one by the above hierarchical procedure. Table I (opposite) shows how the number $n_{r,k}$ of sequences of length r compatible with the constraints for length up to k depends upon r and k.

The way that this table is constructed is by use of a recurrence relation (which I give without proof) for the entries in each column:

Table I.

r \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	...
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	...
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	...
2	4	3	3	3	3	3	3	3	3	3	3	3	3	3	...
3	8	5	4	4	4	4	4	4	4	4	4	4	4	4	...
4	16	8	5	5	5	5	5	5	5	5	5	5	5	5	...
5	32	13	7	7	6	6	6	6	6	6	6	6	6	6	...
6	64	21	9	9	7	7	7	7	7	7	7	7	7	7	...
7	128	34	12	12	8	8	8	8	8	8	8	8	8	8	...
8	256	55	16	16	10	10	10	9	9	9	9	9	9	9	...
9	512	89	21	21	12	12	12	10	10	10	10	10	10	10	...
10	1024	144	28	28	14	14	14	11	11	11	11	11	11	11	...
11	2048	233	37	37	17	17	17	12	12	12	12	12	12	12	...
12	4096	377	49	49	20	20	20	13	13	13	13	13	13	13	...
13	8192	610	65	65	24	24	24	15	15	15	15	15	14	14	...
14	16384	987	86	86	29	29	29	17	17	17	17	17	15	15	...
15	32768	1597	114	114	34	34	34	19	19	19	19	19	16	16	...
16	65536	2584	151	151	41	41	41	21	21	21	21	21	17	17	...
17	131072	4181	200	200	49	49	49	23	23	23	23	23	18	18	...
18	262144	6765	265	265	58	58	58	26	26	26	26	26	19	19	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

$$
\begin{aligned}
n_{r,k} + n_{r,k} &= n_{r+1,k} \ (k = 1,\ r + 1 \geq 1),\\
n_{r,k} + n_{r+1,k} &= n_{r+2,k} \ (k = 2,\ r + 1 \geq 1),\\
n_{r,k} + n_{r+1,k} &= n_{r+3,k} \ (k = 3,\ 4,\ r + 1 \geq 2),\\
n_{r,k} + n_{r+2,k} &= n_{r+5,k} \ (k = 5, 6, 7,\ r + 1 \geq 3),\\
n_{r,k} + n_{r+3,k} &= n_{r+8,k} \ (k = 8, 9, 10, 11, 12,\ r + 1 \geq 5),\\
n_{r,k} + n_{r+5,k} &= n_{r+13,k} \ (k = 13, 14, \ldots, 20,\ r + 1 \geq 8),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}
\tag{5}
$$

The numbers 0, 1, 1, 2, 3, 5, 8, 13, 21, ... appearing here are the Fibonacci numbers. Note also that

$$n_{r,k} = r + 1 \quad \text{if} \quad k \geq r, \tag{6}$$

that

$$n_{r,3} = n_{r,4},\ n_{r,5} = n_{r,6} = n_{r,7},\ n_{r,8} = n_{r,9} = \cdots = n_{r,12},\ \text{etc.} \tag{7}$$

and that $n_{r,1}$ and $n_{r,2}$ are the powers of 2 and the Fibonacci numbers, respectively.

It follows from the above recurrence relations that if any bound is placed on k, then $n_{r,k}$ will increase exponentially with r, whereas for the correct Fibonacci sequences (i.e. $k \geq r$), $n_{r,k}$ is just linear in r (and, in fact, is just $r + 1$). Thus, for large r, the chance of avoiding making a mistake, if k is bounded, becomes vanishingly small. It follows, therefore, that for a two-dimensional (or three-dimensional) tiling of this general five-fold (or icosahedral) quasi-periodic type, the assembly must be highly non-local—unless serious mistakes in the resulting structure are to be permitted.

In my opinion, this raises a number of interesting issues. It may, of course, be the case that the quasi-crystalline substances that have been so far produced do not bear any very close relation to the tiling patterns that I have been discussing. Or perhaps they do, but very large numbers of mistakes are continually being made. My impression, however, is that such mistakes would tend to lead to serious trouble in the continuation, rather than just local "glitches." I do, myself, find the regularity of many of the micrograph pictures to be strikingly impressive, and I am inclined to believe that there is indeed some rather close relationship between quasi-crystalline substances and such tiling patterns. If this turns out to be the case, then we may have to face up to the possibility that there could be something remarkable and non-local about the way that they grow (*cf.* Penrose, 1987). There are essentially non-local aspects to quantum theory. Perhaps these aspects are sometimes being exploited in the way that a quasi-crystal—or even an ordinary crystal—might grow.

Added in proof: A recent preprint, "Growing Perfect Quasicrystals," by G. Y. Onoda, P. J. Steinhardt, D. P. Di Vincenzo, and J. E. S. Socolar, proposes a "local" growth procedure for quasicrystals in which tiles would only stick at places where there is a legal choice when all places where there is *no* choice have been exhausted. This is an interesting suggestion—though it is not actually local in the sense intended here.

References

de Bruijn, N. G. (1981). *Proc. Konink. Ned Akad Wetensch* **A84**, 39–53.

Gardner, M. (1977). Extraordinary nonperiodic tiling that enriches the theory of tiles. *Sci. Amer.* **236**, 110–121.

Grünbaum, B. and Shephard (1987). *Tilings and Patterns* (W. H. Freeman).

MacKay, A. L. (1982). Crystallography and the Penrose pattern. *Physica* **114A**, 609.

Penrose, R. (1974). The rôle of aesthetics in pure and applied mathematical research. *Bull. Inst. Math. Applications* **10**, no. 7/8, 266–271,

Penrose, R. (1979). Pentaplexity. *Mathematical Intelligencer* **2**, 32–38 (from *Eureka*, No. 39).

Penrose, R. (1987). Hermann Weyl, space-time and conformal geometry. *Hermann Weyl 1886–1986* (ed. S. Chandrasekharan). E.T.H. Zurich.

Shechtman, D., Blech, I., Gratias, D. and Cahn, J. W. (1984). Metallic phase with long-range orientational order and no translational symmetry. *Phys Rev. Lett.* **53**, 1951–1953.

189
NON-HAUSDORFF RIEMANN SURFACES AND COMPLEX DYNAMICAL SYSTEMS

Twistor Newsletter **28** (1989). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, Longman, 1990

A distinguished mathematician, who was at a meeting I attended describing this idea, referred to it as 'the product of a deranged mind'. I am not sure whether he was serious—but this kind of mathematical structure seems to have relevance to twistor theory!

 Non-Hausdorff Riemann surfaces and complex dynamical systems. *Twistor Newsletter* **28** (1989). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

§I.6.12 Non-Hausdorff Riemann surfaces and complex dynamical systems *by R. Penrose* (**TN** 28, March 1989)

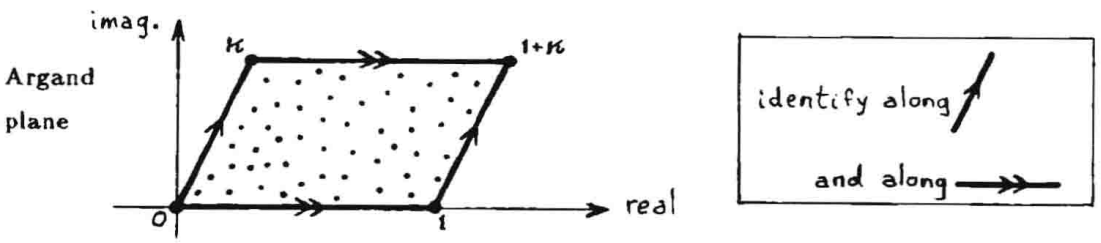
If twistor theory is ever to give a description of such 'non-integrable' systems as the general Yang-Mills or Einstein equations, it must be able to accomodate *chaotic* behaviour (such as occurs in the Belinskii-Lifschitz-Khalatnikov or Misner discussion of Bianchi type 9 vacuums). A simple type of chaotic system (Julia sets, the Mandelbrot set) is given by iterated complex maps

$$z \mapsto z^2 + c \qquad (z,\ c \in \mathbb{C},\ c \text{ constant}).$$

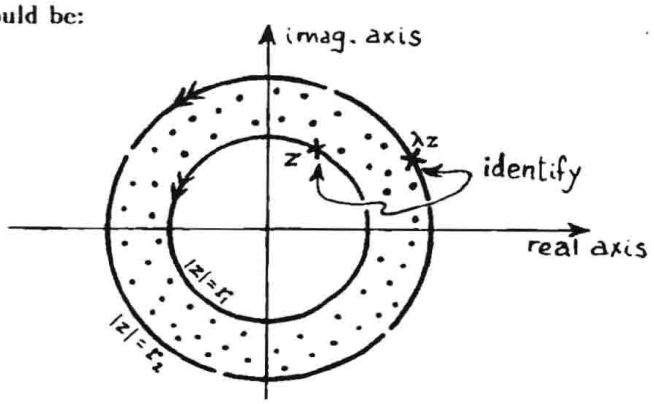
It turns out that such maps have a close relationship to a certain type of non-Hausdorff Riemann surface.

It should be noted that non-Hausdorff Riemann surfaces have already found significant application in the Woodhouse-Mason description of Ward's construction for stationary-axisymmetric solutions of Einstein's vacuum equations, so it is quite possible that non-Hausdorff surfaces of the type that I am considering may also find some significant role within twistor theory.

Let us first recall how an ordinary Riemann surface of genus 1 (torus) may be constructed:



An alternative procedure would be:



Non-Hausdorff Riemann surfaces and complex dynamical systems. *Twistor Newsletter* **28** (1989). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

where the annulus $r_1 \leq |z| \leq r_2$ is curled into a torus by identifying z (at $|z|=r_1$) with λz (at $|z|=r_2$) where $|\lambda|=\frac{r_2}{r_1}$. Thus, in a certain sense, this Riemann surface codes the information of the map

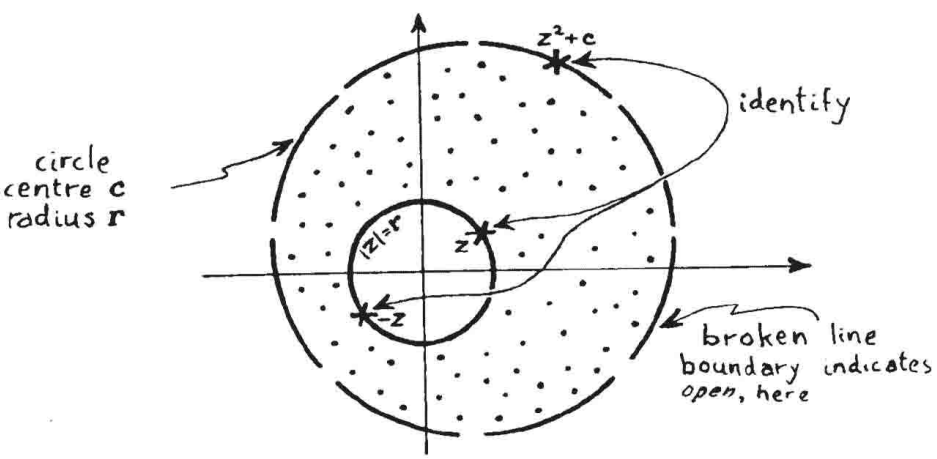
$$z \mapsto \lambda z$$

where we go once around one of the generators of the torus, and iterations of this map correspond to going many times around.

Now consider the *quadratic* map

$$z \mapsto z^2 + c.$$

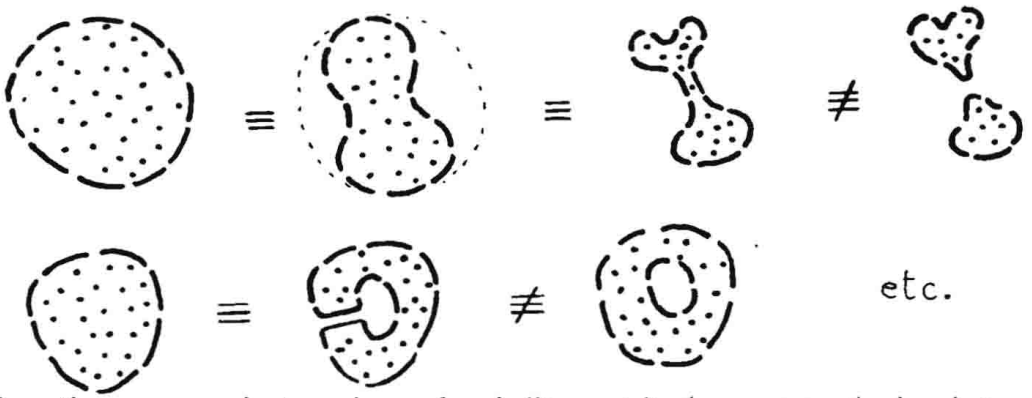
Similarly to the above, we can relate this to a 'Riemann surface' S, constructed thus:



Here the *pair* of opposite points z and $-z$ are to be identified with the *single* point z^2+c. Taking the inner boundary circle to be present and the outer boundary circle to be absent, we have constructed a non-Hausdorff Riemann surface. (Each point has a neighbourhood that is a small complex open disc, but on $|z|=r$, the discs for z and for $-z$ always overlap.)

However, this singles out $|z|=r$ as 'special' which we do not want, so we must adopt a viewpoint that removes this feature. It has already been apparent that some new concept of 'complex manifold' is needed in twistor theory, in which we have an equivalence (for non-compact complex manifolds)

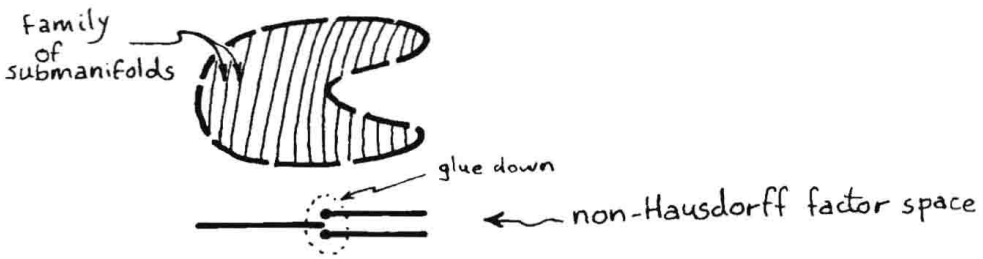
842 Non-Hausdorff Riemann surfaces and complex dynamical systems. *Twistor Newsletter* **28** (1989). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990



Thus, the space can be 'pared away' or built up at its (non-compact) edge, but not so as to give rise to a change in its global structure. The concept remains somewhat vague, as of now, but something of this sort is required if we require elements of

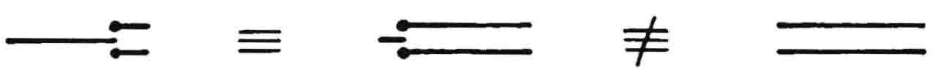
$$H^1(X,\Theta) \qquad (\Theta \text{ is the sheaf of holomorphic vector fields})$$

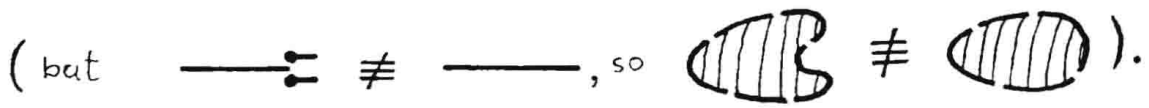
to represent the infinitesimal deformations of a non-compact complex manifold X. (Deformations of the 'boundary' do not show up in $H^1(X,\Theta)$.) Now consider how a non-Hausdorff complex manifold can often arise (as indeed is the case in the Woodhouse-Mason construction):



if we have:

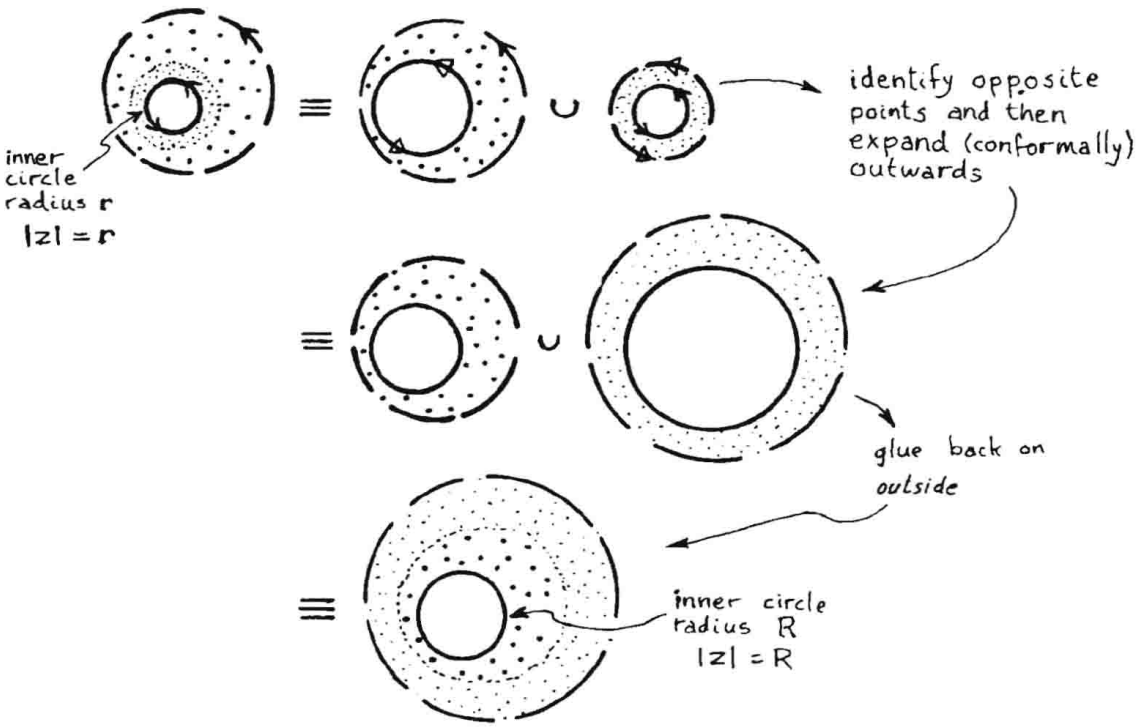
(as above), we should also have:





Thus we can 'split' or 'reglue' non-Hausdorffness, provided that (in some appropriate sense) we do not alter the 'global connectivity' of the space. (I think that this can be stated in terms of what *local* deformations are allowed and what are not—but this needs more understanding.)

In the case of the non-Hausdorff Riemann surface just constructed, we can remove a small annulus from the inner circle boundary and glue it on at the outer circle boundary, but we must identify opposite points of this annulus as we do so:


inner circle radius r
|z| = r
identify opposite points and then expand (conformally) outwards
glue back on outside
inner circle radius R
|z| = R

Also we can perform this operation in reverse so long as there remains room on the inside.

If c is in the Mandelbrot set M, then the topological structure of S (under this equivalence) appears to differ in an essential way from when c is not in M. For in some sense S has a 'covering space' which is the complement K of the Julia set in $\mathbf{C}$, and whether c is in M or not depends on the multiple-connectivity properties of K. (When $c \notin M$, K has a Cantor set removed from it, but when $c \in M$, K is topologically an annulus.)

844 Non-Hausdorff Riemann surfaces and complex dynamical systems. *Twistor Newsletter* **28** (1989). Also published in *Further Advances in Twistor Theory*, Vol. 1. Eds L. J. Mason and L. P. Hughston, 1990

Some clarifying ideas seem to be needed.

References

Belinskii, V.A., Khalatnikov, I.M. & Lifschitz, E.M. (1970) Oscillatory approach to a singular point in the relativisic cosmology, Adv. Phys. **19**, 523-573.

Misner, C.W. (1969) Mixmaster universe, Phys. Rev. Lett., 22, 1071-4.

Woodhouse, N.M.J. & Mason, L.J. (1988) The Geroch group and non-Hausdorff twistor spaces, Nonlinearity **1**, 73-114.

190
USING SPINORS IN GENERAL RELATIVITY

Proc. Work. Gravitation, Magneto-Convection and Accretion (in honour of the 60th birthdays of Friedrich Meyer and Jürgen Ehlers, Ringberg Castle, Tegernsee, May 28–31, 1989). Eds Bernd Schmidt, Hermann-Ulrich Schmidt, and Hans-Christoph Thomas, 56–60. Max-Planck Institut für Physik und Astrophysik, 1989

846 Using spinors in general relativity. *Proc. Work. Gravitation, Magneto-Convection and Accretion* (in honour of the 60th birthdays of Friedrich Meyer and Jürgen Ehlers, Ringberg Castle, Tegernsee, May 28–31, 1989). Eds Bernd Schmidt, Hermann-Ulrich Schmidt, and Hans-Christoph Thomas, 56–60. Max-Planck Institut für Physik und Astrophysik, 1989

Using Spinors in General Relativity

Roger Penrose
Mathematical Institute
Oxford, U.K.

The role of spinors in the standard quantum-mechanical description of particles of non-integral spin is well known. Somewhat less well known is that spinors provide a very useful calculus for the description of space-time geometry itself, in some ways more basic and elementary than the tensor calculus normally employed in descriptions of (special or) general relativity. To be at its most effective, it is necessary that the two-component form of the spinor calculus be used (with essentially the notation due to van der Waerden[1]), rather than the four-component ("Dirac") form that is rather more familiar to physicists[2].

It would be inappropriate for me to attempt to review all the various developments in applications of spinors to general relativity, but I shall mention a few important examples. In particular, spinors can be used to study ray optics in space-time, and provide neat descriptions of space-time curvature. Some important early work in this area was done by Jürgen Ehlers, in his well-known collaboration with Ray Sachs and Pascual Jordan[3]. It was only very recently that I became aware that the spinor parts of this paper were due entirely to Jürgen. Thus, it was this spinor work that represented the very beginning of that long list of important contributions that Jürgen has made to general relativity. (The spinor components that enter into these descriptions are examples of "spin-coefficients", which form the basis of what has become known as the Newman-Penrose formalism.[4]) It is thus most appropriate that I can pay my respects to Jürgen by describing certain applications of spinor techniques to general relativity. I shall also indicate certain new ideas that I have been working with recently, which could give a different angle on ways that spinors can be applied in general relativity.

First, let me remind people of one of the earliest and most striking simplifications that 2-spinors provide, namely in the classification of vacuum Riemann tensors (or of Weyl tensors, if matter is present). We should recall, first of all, that a 2-spinor ζ^A has a geometrical realization (up to sign) as a null flag: a future-null vector (a vector, called the "flagpole", pointing up the light cone) together with a two-dimensional null half-plane ("the flag plane") attached. The translation from tensor to spinor components can be achieved via the basic relation

$$\begin{pmatrix} v^{00'} & v^{01'} \\ v^{10'} & v^{11'} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} v^0 + v^3 & v^1 + iv^2 \\ v^1 - iv^2 & v^0 - v^3 \end{pmatrix}$$

which shows how a tensor quantity v^a, with a single tensor index in upper position, can be translated to a corresponding spinor quantity $v^{AA'}$, with a corresponding pair of spinor indices, one unprimed and one primed. (Primed indices refer to the complex conjugate spin space.) This translation appropriately extends to tensors with any number of indices.

Using spinors in general relativity. *Proc. Work. Gravitation, Magneto-Convection and Accretion* (in honour of the 60th birthdays of Friedrich Meyer and Jürgen Ehlers, Ringberg Castle, Tegernsee, May 28–31, 1989). Eds Bernd Schmidt, Hermann-Ulrich Schmidt, and Hans-Christoph Thomas, 56–60. Max-Planck Institut für Physik und Astrophysik, 1989

and to indices in the lower position as well as upper. With this tensor-spinor translation scheme we can rewrite the vacuum Riemann curvature tensor R_{abcd}, or more generally the Weyl tensor C_{abcd} if we allow matter to be present, as

$$C_{AA'BB'CC'DD'} = \Psi_{ABCD}\epsilon_{A'B'}\epsilon_{C'D'} + \epsilon_{AB}\epsilon_{CD}\bar{\Psi}_{A'B'C'D'} \quad ,$$

with Ψ_{ABCD} totally symmetrical. We now find that Ψ_{ABCD} can be decomposed as a symmetrized product of single-index spinors:

$$\Psi_{ABCD} = \alpha_{(A}\beta_B\gamma_C\delta_{D)} \quad .$$

Each spinor α^A, β^A, γ^A, δ^A defines a "flagpole" direction along the light cone in accordance with the above descriptions. (The flag plane directions are not significant here because the individual spinors are fixed only up to proportionality.) These four flagpole directions are called the gravitational principal null directions, and their possible coincidence scheme determine the required classification[5,6,2] of C_{abcd}.

As a simple application of this classification, we can very directly deduce the fact that all Lemaitre-Friedmann-Robertson-Walker models are conformally flat. We can see this immediately because the spatial isotropy of these models implies that the gravitational principal null directions must vanish (otherwise their spatial projections would not be invariant under spatial rotation) whence $\Psi_{ABCD} = 0$, i.e. $C_{abcd} = 0$, which is indeed the condition for conformal flatness.

An important more recent application of spinor techniques is that introduced by Witten[7], to give a proof of the positivity of total energy in general relativity, where it is assumed that the local energy-momentum tensor of matter satisfies the so-called "dominant energy condition". Spinors also play a key role in providing plausible "quasi-local" expressions for energy(-momentum/angular momentum) which refer not to the entire total energy (etc.), but to the total that threads through a finite spacelike 2-surface[8,2]. It is remarkable that spinors, which are basically "spin 1/2" objects, should provide such fundamentally new insights, in a general-relativistic context, into the concepts of energy and momentum, which had previously been thought of as objects described by quantities of just integral spin.

The spinors that are used in these descriptions are the 2-spinors that are referred to above, in which tensor indices (lower case) are translated into spinor form (capital indices) according to a prescription

$$a \leftrightarrow AA', \quad b \leftrightarrow BB', \quad c \leftrightarrow CC', \ldots, a_0 \leftrightarrow A_0A_0' \ldots, \quad a_1 \leftrightarrow A_1A_1' \ldots .$$

the space-time being four-dimensional and with Lorentzian signature -2. (The choice $+---$, giving the signature -2 fits in rather better with the spinor notation than does $-+++$, giving signature $+2$, but the distinctions are not very important for us.) Spinors are also important in spaces of other dimensions and signatures, and sometimes such spinors can be applied in slightly indirect ways to obtain results for ordinary four-dimensional Lorentzian space-time, as we shall see in a moment.

848 Using spinors in general relativity. *Proc. Work. Gravitation, Magneto-Convection and Accretion* (in honour of the 60th birthdays of Friedrich Meyer and Jürgen Ehlers, Ringberg Castle, Tegernsee, May 28–31, 1989). Eds Bernd Schmidt, Hermann-Ulrich Schmidt, and Hans-Christoph Thomas, 56–60. Max-Planck Institut für Physik und Astrophysik, 1989

Let us consider the proper rotation group $SO(p,q)$, i.e. the group of (pseudo-)rotations of unit determinant perserving the quadratic form

$$x_1^2 + x_2^2 + \ldots + x_p^2 - y_1^2 - y_2^2 - \ldots - y_q^2$$

The dimension is $n = p+q$ and the signature $s = p-q$. When n is even, the unreduced spin-space is $2^{n/2}$-dimensional, but it splits into a direct sum of two spaces, each of dimension $2^{n/2-1}$, the "unprimed" and "primed" spaces, respectively. (When $n = 4$, these reduced spaces are the familiar two-dimensional spin-spaces of ζ^A and $\eta^{A'}$, respectively, while the unreduces spin-space is the four-dimensional space of Dirac spinors, namely the direct sum of these two.) When n is odd, the spin space does not reduce in this way, but is a single space of dimension $2^{(n-1)/2}$. As in the case $n = 4$, I shall use unprimed and primed capital indices for the two reduced spin spaces whenever n is even, and lower case Greek indices when n is odd.

In addition to this rapid (exponential) increase with n in the dimensionality of spin-space (which tends to imply that spinors are comparatively more useful when n is reasonably small), there is a curious periodicity, with period 8, in the way that spinors are related to the corresponding (n-dimensional) vectors[9].

n (mod 8):	0	1	2	3	4	5	6	7
index transl.:	AB'	$(\alpha\beta)$	(AB)	$(\alpha\beta)$	AB'	$[\alpha\beta]$	$[AB]$	$[\alpha\beta]$
ε :	(AB)	$(\alpha\beta)$	AB'	$[\alpha\beta]$	$[AB]$	$[\alpha\beta]$	AB'	$(\alpha\beta)$

Here, the second line provides the way that tensor indices are to be translated as pairs of spinor indices, round brackets referring to symmetry and square brackets to antisymmetry. (When a pair AB appears, with neither index being primed, then the primed pair $A'B'$ would do equally well.) The third line refers to the (symmetry of the) "epsilon spinor". (When it has the form $\varepsilon_{AB'}$ or $\varepsilon^{AB'}$ it can be used to eliminate primed indices by converting them into unprimed ones in the opposite position.) There is also a periodicity mod 8 in the signature s:

s (mod 8)	0	1	2	3	4	5	6	7
	R	**R**	$\mathbb{C}$	**H**	**H**	**H**	$\mathbb{C}$	**R** ,

where **R**, $\mathbb{C}$, **H** refer the number systems (real, complex, quaternionic) that apply to the spinors when the corresponding vectors are real.

One particular case that has considerable interest in relativity occurs when $n = 6$ and $s = -2$ ($SO(2,4)$ being essentially the conformal group of Minkowski space). The reduced spinors are now what are called twistors[2]. The most direct way in which twistors can be used to describe space-time geometry is in the representation of light rays (null geodesics) in Minkowski space. A symbol Z^α is often used to denote a twistor, where I am now using Greek letters to denote the four-dimensional reduced spinors, even though the space is even-dimensional, in order to avoid confusion with the ordinary space-time 2-spinor indices A, A', etc. We can write

$$Z^\alpha = (\omega^A, \pi_{A'})$$

giving Z^α in terms of its two "spinor parts". The complex conjugate $\overline{Z}_\alpha$ of Z^α is given by

$$\overline{Z}_\alpha = (\overline{\pi}_A, \overline{\omega}^{A'}) \quad ,$$

and we call Z^α a null twistor if $Z^\alpha \overline{Z}_\alpha (= \omega^A \overline{\pi}_A + \pi_{A'} \overline{\omega}^{A'}) = 0$. Null twistors (up to proportionality) describe light rays in Minkowski space (possibly at infinity). The flagpole of $\pi_{A'}$ determines the direction of the ray, and if x^a is the position vector of a point on the ray, then

$$\omega^A = ix^{AA'} \pi_{A'} \quad .$$

In curved space-time, such a direct and global description of a twistor does not arise. However, one can use what are called hypersurface twistors,[2] which refer to a choice of a (normally spacelike and analytic) hypersurface H in the space-time. Such twistors can be used in the study of the initial value problem in general relativity, and there are exciting possibilities that this approach may relate closely to Ashtekar's important "new variables". However, as yet progress has been somewhat hampered in this approach owing to the awkwardness of the mathematical structure that arises for such hypersurface twistor spaces. This structure is known as a CR-structure and the most effective way of handling it[10] seems to be by introducing what is known as a Fefferman metric. In the case that arises here, the Fefferman metric turns out to be a quadratic form for which $n = 6$ and $s = 0$. It seems likely that the spinors for this Fefferman metric will be of considerable value.

There is also the strong possibility that the "twistors" for the Fefferman metric could have significance. These would be reduced spinors for the case $n = 8$, $s = 0$ ("second twistorization"!). There is a special interest in the case $n = 8$, especially when $s = 0$, 8 or -8, because there is what is known as a "principle of triality", in which the unprimed spinors, the primed spinors and the original vectors are all on an equal footing with one another.

As an initial step towards developing a "spinor" or "twistor" calculus for this Fefferman metric, I have recently been examining the "8-twistors" that arise in this context in the case of Minkowski space. These would be the reduced spinors for the quadratic form

$$Z^\alpha \overline{Z}_\alpha \equiv \omega^A \overline{\pi}_A + \pi_{A'} \overline{\omega}^{A'} = 0$$

(which indeed has $n = 8$, $s = 0$). Now the "vectors" are to be the original null twistors Z^α (or, rather their real and imaginary parts – or, more appropriately, the pairs Z^α, $\overline{Z}_\alpha$, with $Z^\alpha \overline{Z}_\alpha = 0$). One finds[11] that the "unprimed spinors" are defined by pairs $(y^{AA'}, \lambda)$ where y^a is real and λ is complex, and where there is an "incidence relation" with $Z^\alpha = (\omega^A, \pi_{A'})$, given by

$$\omega^A - iy^{AA'} \pi_{A'} = \lambda \overline{\pi}^A \quad .$$

Curiously, the "primed spinors" are null twistors again, now Y^α (or the pairs Y^α, $\overline{Y}_\alpha$, with $Y^\alpha \overline{Y}_\alpha = 0$), with a similar incidence relation with $(y^{AA'}, \lambda)$ as Z^α had, and where the incidence between Z^α and Y^α is expressed by

$$Z^\alpha Y^\beta - Y^\alpha Z^\beta = \epsilon^{\alpha\beta\gamma\delta} \overline{Z}_\gamma \overline{Y}_\delta \quad .$$

850 Using spinors in general relativity. *Proc. Work. Gravitation, Magneto-Convection and Accretion* (in honour of the 60th birthdays of Friedrich Meyer and Jürgen Ehlers, Ringberg Castle, Tegernsee, May 28–31, 1989). Eds Bernd Schmidt, Hermann-Ulrich Schmidt, and Hans-Christoph Thomas, 56–60. Max-Planck Institut für Physik und Astrophysik, 1989

References

[1] Waerden, B.L. van der (1929) Spinoranalyse, Nachr. Akad. Wiss. Göttingen Math. Physik. Kl, 100–109

[2] For a detailed account of 2-spinor calculus, see Penrose, R. and Rindler, W. (1984) Spinors and Space-Time, Vol. 1: Two-Spinor Calculus and Relativistic Fields and (1986) Vol. 2: Spinor and Twistor Methods in Space-Time Geometry (Cambridge University Press, Cambridge).

[3] Jordan, P., Ehlers, J. and Sachs, R.K. (1961). Beiträge zur Theorie der reinen Gravitationsstrahlung, Akad. Wiss. Lit. Mainz, Abhandl. Math. Nat. Kl, 1960 Nr. 1

[4] Newman, E.T. and Penrose, R. (1962) An approach to gravitational radiation by a method of spin coefficients. J. Math. Phys. 3, 566–78; Errata: ditto ibid. 4, 998.

[5] Cartan, E. (1922) Sur les équations de la gravitation d'Einstein, I. Math. Pur. et Appl. 1, 141–203, (cf.p.194).

[6] Penrose, R. (1960) A spinor approach to general relativity, Ann. Phys. 10, 171–201.

[7] Witten, E. (1981) A new proof of the positive energy theorem Comm. Math. Phys. 80, 381–402.

[8] Penrose, R. (1982) Quasi-local mass and angular momentum in general relativity, Proc. Roy. Soc. London A381, 53–63.

[9] See appendix to ref.2, vol.2

[10] A result due to G.A.J. Sparling.

[11] Some closely related work, obtaining very similar results had earlier been reported in a Twistor Newsletter article by L.P. Hughston.

191
A TWISTORIAL APPROACH TO THE FULL VACUUM EQUATIONS

Twistor Newsletter **29** (1989). With L. J. Mason. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. Chapman and Hall/CRC, London, 2001

 A twistorial approach to the full vacuum equations. *Twistor Newsletter* **29** (1989). With L. J. Mason. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

§III.3.19 A Twistorial Approach to the Full Vacuum Equations *by L.J. Mason and R. Penrose* (TN 29, November 1989)

In an attempt to explore possible connections between Ashtekar variables and hypersurface twistors, we have been led to a (tentative) new approach to the twistorial description of general relativity.

Let $\mathcal{M}$ be a complexified (analytic) space-time, and $\mathscr{H}$, a complexified (analytic) spacelike hypersurface in $\mathcal{M}$. A point in the projective *hypersurface twistor space* $\mathbb{P}\mathcal{J} = \mathbb{P}\mathcal{J}(\mathscr{H})$ is represented as an *α-curve* in $\mathscr{H}$. An α-curve is a (complex) null curve in $\mathscr{H}$ with tangent vector of the form $\tilde{\pi}^A \pi^{A'}$, where

$$\tilde{\pi}_A = t_A{}^{A'} \pi_{A'},$$

the vector t^a being *normal* to $\mathscr{H}$ (and can be normalized, *if so desired*, by

$$t_a t^a = 2, \quad \text{so that} \quad t_A{}^{B'} t_{B'}{}^C = -\varepsilon_A{}^C \quad \text{and} \quad t_{A'}{}^B t_B{}^{C'} = -\varepsilon_{A'}{}^{C'}$$

whence $\tilde{\tilde{\pi}}_{A'} = -\pi_A$, with $\tilde{\lambda}_{A'} = t_{A'}{}^A \lambda_A$ etc.), and where we *require*

$$\tilde{\pi}^A \pi^{A'} \nabla_{AA'} \pi_{B'} = 0 \tag{A}$$

The *non*-projective hypersurface twistor space $\mathcal{J} = \mathcal{J}(\mathscr{H})$ is a line bundle over $\mathbb{P}\mathcal{J}$, with fibres given by the scalings for $\pi_{A'}$. (Penrose and Rindler 1986 pp. 215-218).

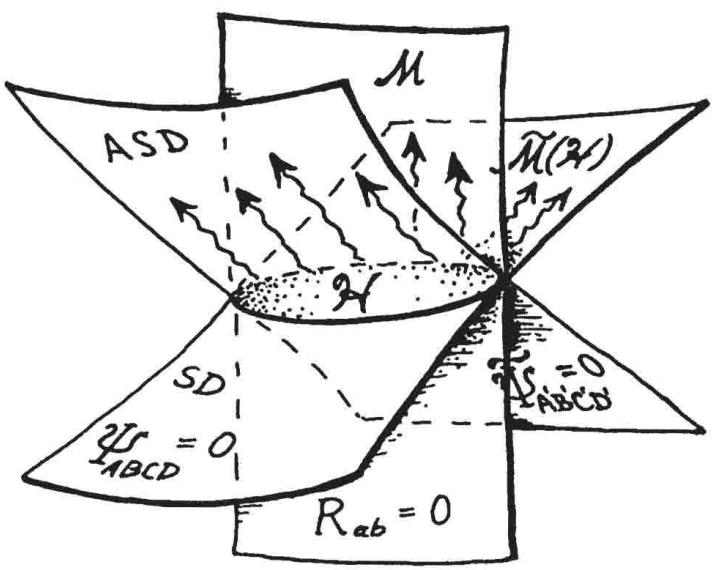
As a point of contact with Ashtekar's variables, note that the $\nabla_{AA'}$ in (A) is his (primed) *spin connection* — though only the *conformal* part of it is being used here — and his *soldering form* is implicitly involved in $\tilde{\pi}^A \pi^{A'} \nabla_{AA'}$. To obtain the *entire* (primed) spin connection we should need to incorporate

$$\tilde{\pi}^A \pi^{A'} \nabla_{AA'} \nu_{B'} \quad \text{with } \nu_{B'} \textit{ independent of } \pi_{B'} \tag{B}$$

A helpful concept, in relation to hypersurface twistors is that of "Heaven on Earth" $\mathcal{M}(\mathscr{H})$, which is the "nonlinear graviton" construction applied to $\mathbb{P}\mathcal{J}$ (or $\mathcal{J}$); i.e., we take holomorphic curves in $\mathbb{P}\mathcal{J}$ (homologous to those representing points in $\mathscr{H}$) to represent points in $\mathcal{M}(\mathscr{H})$. Note that $\mathcal{M}(\mathscr{H})$ is conformally anti-self-dual (but not conformal to vacuum, in general), and we may regard

$$\mathscr{H} \subset \mathcal{M}(\mathscr{H}).$$

(We could define $\tilde{\mathcal{M}}(\mathscr{H})$ similarly, as a self-dual space, where primed and unprimed spinor indices are interchanged.) The picture below



may be helpful. If $\mathcal{M}$ is vacuum, we can regard $\mathcal{M}$ as the Einstein evolution of $\mathscr{H}$, whereas $\mathcal{M}(\mathscr{H})$ is a conformal manifold evolved from the (conformally invariant part

of) the same data at $\mathscr{H}$, but evolved according to $\tilde{\Psi}_{A'B'C'D'} = 0$. The space $\mathcal{M}(\mathscr{H})$ is "equivalent information" to $\mathcal{J}$, and it provides a useful way of studying the global structure of $\mathcal{J}$. (Similarly, the conformally self-dual space $\check{\mathcal{M}}(\mathscr{H})$, evolved from $\mathscr{H}$ according to $\Psi_{ABCD} = 0$, gives equivalent information to the "dual" hypersurface twistor space defined by β-curves (complex conjugation of (A) on $\mathscr{H}$).

The complex conformal space $\mathcal{M}(\mathscr{H})$ provides a good part of the initial data needed at $\mathscr{H}$ for the evolution of Einstein's equations (to give $\mathcal{M}$). However, two-additional ingredients are needed:

1. a conformal scale (equivalent to (B) above)
2. the location of $\mathscr{H}$ itself, within $\mathcal{M}(\mathscr{H})$.

Here we make use of an idea (cf. §III.3.18 and LeBrun §III.1.4) whereby 1. and 2. are coded as two solutions of the conformally invariant wave equation on $\mathcal{M}(\mathscr{H})$

$$(\Box + \frac{R}{\sigma})\Omega = 0, \quad (\Box + \frac{R}{\sigma})\chi = 0.$$

Using $g_{ab} \mapsto \Omega^2 g_{ab}$, we scale $\Omega \mapsto \Omega^{-1}\Omega = 1$ (so that R is scaled to zero) and $\chi \mapsto \Omega^{-1}\chi$, and we use $\chi = 0$ to define the location of $\mathscr{H}$. The freedom in Ω is precisely the freedom in choosing a conformal scale on $\mathcal{M}(\mathscr{H})$ so that $R = 0$; the freedom in χ is the freedom in the choice of $\mathscr{H}$ within $\mathcal{M}(\mathscr{H})$ together with the choice of "lapse function", provided by the choice of normal (where now t_a need not be normalized by $t_a t^a = 2$)

$$t_a = \nabla_a \chi \quad \text{at } \mathscr{H}$$

in $\mathcal{M}(\mathscr{H})$ — or, equivalently, in $\mathcal{M}$. (Note that $\chi = 0$ at $\mathscr{H}$, together with $\nabla_a \chi$, give initial data for $(\Box + \frac{R}{\sigma})\chi = 0$ in $\mathcal{M}(\mathscr{H})$.)

We could specify Ω and χ as two elements of $H'(\mathcal{O}(-2))$ on $\mathbb{P}\mathcal{J}$ — but we prefer to use a *bundle* $\mathcal{B}$ over $\mathbb{P}\mathcal{J}$ (or $\mathcal{J}$). We take $\mathcal{B}$ as an affine rank-2 bundle whose fibres are solution of

$$\nabla_A \nu_{A'} = \nabla_{AA'}\chi \tag{C}$$

on an α-plane in $\mathcal{M}(\mathscr{H})$. (Note that, being ASD, $\mathcal{M}(\mathscr{H})$ has α-planes which meet $\mathscr{H}$ in the α-curves — and each α-plane corresponds to a point of $\mathbb{P}\mathcal{J}$.) Here we define

$$\nabla_A = \pi^{A'} \nabla_{AA'}$$

($\nabla_{AA'}$ is the derivative operator in $\mathcal{M}(\mathscr{H})$) which differentiates within the α-plane, and we assign a homogeneity -1 to $\nu_{A'}$, with respect to $\pi_{A'}$. (Thus, under $\pi_{A'} \mapsto \zeta\pi_{A'}$, we take $\nu_{A'} \mapsto \zeta^{-1}\nu_{A'}$.) Note that χ has homogeneity 0. Observe that since $\tilde{\Psi}_{A'B'C'D'} = 0$ in $\mathcal{M}(\mathscr{H})$, we have

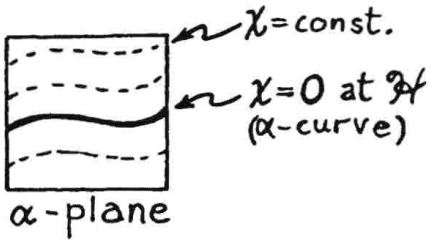
$$\nabla_A \nabla_B \nu_{C'} - \nabla_B \nabla_A \nu_{C'} = 0$$

which demands $\Box\chi = 0$ and which we have taken to be so. We note that

$$X[\nu] = \chi - \nu_{A'}\pi^{A'} \in \mathbb{C}$$

is constant on the α-plane, so we have a bundle map

$$X : \mathcal{B} \to \mathcal{J} \times \mathbb{C}.$$



This bundle map leads to the exact sequence of affine bundles

$$0 \to \mathcal{B}_0 \longrightarrow \mathcal{B} \xrightarrow{X} \mathcal{O} \to 0. \qquad \text{(D)}$$

The affine 1-bundle $\mathcal{B}_0$ is the zero set of X and encodes the cohomology class of χ alone, and the vector bundle $\hat{\mathcal{B}}$, the underlying translation bundle, or bundle of displacements of $\mathcal{B}$, encodes Ω alone.

The line bundle $\mathcal{B}_0$ has $\mathcal{O}(-2)$ as its bundle of displacements, so that (D) leads to an exact sequence of vector bundles

$$0 \to \mathcal{O}(-2) \to \hat{\mathcal{B}} \to \mathcal{O} \to 0, \qquad \text{(E)}$$

where points of the fibre of $\hat{\mathcal{B}}$ are differences between solutions of (C), i.e., solutions of $\nabla_A \nu_{A'} = 0$ on $\mathcal{M}(\mathscr{H})$. This codes the needed information of the primed spin connection. The cohomology class of $\Omega \in H^1(\mathcal{O}(-2))$ is the extension class that determines the sequence (E).

The cohomology class of $\chi \in H^1(\mathcal{O}(-2))$ determines $\mathcal{B}_0$ (see Eastwood §I.2.9). Geometrically $\mathcal{B}_0$ encodes the location of $\mathscr{H}$ in $\mathcal{M}(\mathscr{H})$ as follows. If $\mathcal{L}$ is a holomorphic curve in $\mathbb{P}\mathcal{J}$ corresponding to a point L in $\mathcal{M}(\mathscr{H})$, a section $\mathcal{S}$ of $\mathcal{B}$ over $\mathcal{L}$ is a $\nu_{A'}(\pi_{B'})$ at L with homogeneity -1 in $\pi_{B'}$. The spinor $\nu_{A'}$ must therefore vanish and so X will map all of $\mathcal{S}$ to the value of χ in $\mathbb{C}$, which is 0 iff $L \in \mathscr{H}$. The total space of $\mathcal{B}$ can be canonically identified with the spin bundle of $\mathcal{M}(\mathscr{H})$ and $\mathcal{B}_0$, with its restriction to $\mathscr{H}$.

How does Einstein's equation fit into all this? How do we formulate the constraint equations and evolution equations? The following curious fact may be useful. Let $(\rho^A, \sigma_{A'})$ be a *local twistor* which is *constant* on the α-plane (consistent because $\tilde{\Psi}_{A'B'C'D'} = 0$). We have

$$\left.\begin{aligned} \nabla_A \rho^B &= -\mathrm{i}\,\pi^{A'} \sigma_{A'} \varepsilon_A{}^B \\ \nabla_A \sigma_{B'} &= -\mathrm{i}\,P_{ABA'B'} \pi^{A'} \rho^B \end{aligned}\right\} \quad \text{on an } \alpha\text{-plane in } \mathcal{M}(\mathcal{H})$$

(Note that $P_{ABA'B'} = \Phi_{ABA'B'} = P_{(AB)(A'B')}$ on $\mathcal{M}(\mathscr{H})$.) These equations of local twistor transport are the Ward transform of $T^*(-1)\mathcal{J}$, see LeBrun 1986 and Woodhouse 1985. Consider $\tilde{\sigma}_A = t_A{}^{A'} \sigma_{A'}$ (with t_a not necessarily normalized); then we have

$$\nabla^A \tilde{\sigma}_A = 0 \text{ at } \chi = 0 \text{ on } \mathcal{M}(\mathscr{H}) \qquad (F)$$

by virtue of $\Box \chi = 0$ and of

$$P_{ABA'B'}\, t^{AA'} = 0 \text{ at } \chi = 0 \text{ on } \mathcal{M}(\mathscr{H}). \qquad (G)$$

Equation (G) states the *constraint equations* for Einstein's theory, the $P_{ABA'B'}$ for $\mathcal{M}(\mathscr{H})$, at $\mathscr{H}$, being a linear combination of

$$\overset{\bullet}{\Psi}_{ABCD} t^C_{A'} t^D_{B'}, \quad \overset{\bullet}{\tilde{\Psi}}_{A'B'C'D'} t^{C'}_A t^{D'}_B, \quad \overset{\bullet}{\Phi}_{ABA'B'}, \quad \overset{\bullet}{\Lambda} t_{AA'} t_{BB'}$$

where the curvature quantities with dots $\bullet$ beneath them denote the *physical* quantities for $\mathcal{M}$. (Before, we had been using the quantities for $\mathcal{M}(\mathscr{H})$.) The terms in $\overset{\bullet}{\Psi}_{ABCD}$ and $\overset{\bullet}{\tilde{\Psi}}_{A'B'C'D'}$ disappear, in $P_{ABA'B'} t^{AA'}$, and we are left only with the Ricci tensor parts (because of the symmetry of $\overset{\bullet}{\Psi}_{ABCD}$, $\overset{\bullet}{\tilde{\Psi}}_{A'B'C'D'}$). Thus, by (F)

$$\tilde{\sigma}_A = \nabla_A \xi \quad \text{on } \chi = 0 \quad (\text{in } \mathcal{M}(\mathscr{H})), \text{ for some } \xi$$

states the constraint equations for the Einstein theory!

What is ξ? We can find a "potential" for $P_{ABA'B'}$ at $\chi = 0$ (on an α-plane in $\mathcal{M}(\mathscr{H})$) by solving

$$\{\nabla_A \nabla_B + \pi^{A'} \pi^{B'} P_{ABA'B'}\} \eta = P_{ABA'B'}\, \nu^{A'} \pi^{B'}$$

and then we find that we can put

$$\xi = \sigma_{A'}(\nu^{A'} - \eta \pi^{A'}) + \rho^A (\mathrm{i}\,\nabla_A \eta) = \sigma_{A'} \mu^{A'} + \rho^A \lambda_A = S^\alpha W_\alpha$$

where

$$\begin{cases} S^\alpha = (\rho^A, \sigma_{A'}) \\ W_\alpha = (\mathrm{i}\nabla_A \eta, \nu^{A'} - \eta\pi^{A'}) = (\lambda_A, \mu^{A'}) \end{cases}$$

on the α-plane in $\mathcal{M}(\mathscr{H})$. Here W_α is *not* local twistor constant on the α-plane because of the term $\nabla_A^{B'}\chi$ in

$$\left.\begin{aligned} \nabla_A \mu^{B'} &= \mathrm{i}\lambda_A \pi^{B'} + \nabla_A^{B'}\chi \\ \nabla_A \lambda_B &= \mathrm{i} P_{ABA'B'} \mu^{A'} \pi^{B'} \end{aligned}\right\} \quad \begin{array}{l}\text{By Einstein (constraint eqns.): integrable} \\ \text{to first order away from } \chi = 0.\end{array}$$

This suggests that we do the following: instead of using $\mathcal{B}$, why not use $\mathcal{A}$, where the fibre over a point Z^α of $\mathcal{J}$ is a $W_\alpha = (\lambda_A, \mu^{A'})$ satisfying these two equations (or rather, these equations contracted with $\tilde{\pi}^A$). This is an "affinized" version of the cotangent bundle $T^*(-1)\mathcal{J}$, of $\mathcal{J}$, referred to above, and we can construct the bundle of displacements of $\hat{\mathcal{A}}$ of $\mathcal{A}$ (as with $\hat{\mathcal{B}}$ above) so that $\hat{\mathcal{A}} = T^*(-1)\mathcal{J}$. (Thus the $\nabla_A^{B'}\chi$ is removed in the passage to $\hat{\mathcal{A}}$, and $\hat{\mathcal{A}}$ *is* the cotangent bundle.) As with $\mathcal{B}$, we have a bundle map

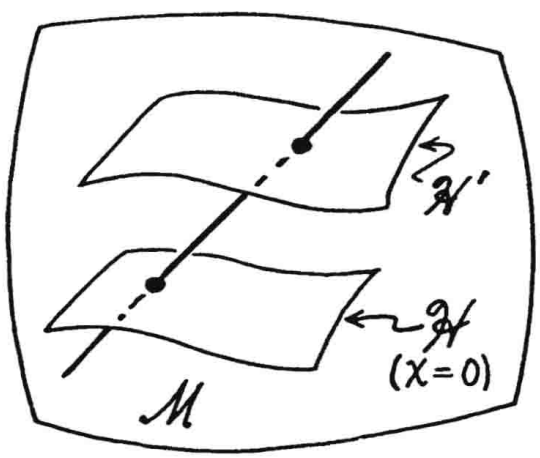
$$X : \mathcal{A} \to \mathcal{J} \times \mathbb{C}$$

where

$$X(W_\alpha) = \chi + \mu^{A'}\pi_{A'} = \chi + W_\alpha Z^\alpha \qquad (Z^\alpha \text{ represented by } (0, \pi_{A'}) \text{ on } \alpha\text{-plane})$$

(On $\hat{\mathcal{A}}$ this gives $X = W_\alpha Z^\alpha$.) As with $\mathcal{B}$, $\mathcal{A}$ serves to define $\mathscr{H}$ within $\mathcal{M}(\mathscr{H})$, and we believe that $\mathcal{A}$ fixes Ω also (not yet checked!).

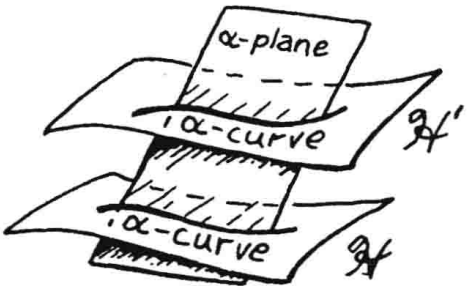
What we should *like* to do would be to *evolve* the Einstein equations and obtain $\mathcal{M}$.



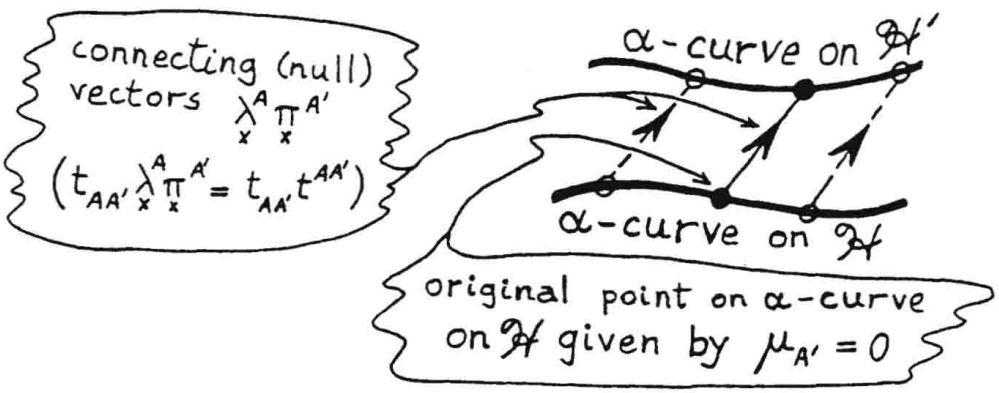
This would require something like using null geodesics defined by $W_\alpha Z^\alpha = 0$ (ambitwistors) to map from one $\mathscr{H}$ to another $\mathscr{H}'$ and to extend *away* from $W_\alpha Z^\alpha = 0$

 A twistorial approach to the full vacuum equations. *Twistor Newsletter* **29** (1989). With L. J. Mason. Also published in *Further Advances in Twistor Theory*, Vol. 3. Eds L. J. Mason, L. Hughston, P. Z. Kobak, and K. Pulverer. 2001

(preferably to all orders, in some way). This would require our being able, given (W_α, Z^α) on $\mathscr{H}$, to find (W_α, Z^α) on $\mathscr{H}'$. Take $\mathscr{H}'$ to be infinitesimally above $\mathscr{H}$. What we need is a way of moving an α-curve from $\mathscr{H}$ to $\mathscr{H}'$, given a W_α on it.



In the *flat* case, this movement is just that given by the α-plane in $\mathcal{M}$ (which now *exists*). The displacement is given by


connecting (null) vectors $\underset{x}{\lambda}^A \underset{x}{\pi}^{A'}$
$(t_{AA'} \underset{x}{\lambda}^A \underset{x}{\pi}^{A'} = t_{AA'} t^{AA'})$
α-curve on
α-curve on
original point on α-curve on given by $\mu_{A'} = 0$

If $W_\alpha Z^\alpha = 0$, we could start off the null geodesic defined by (W_α, Z^α) at the point on the α-curve on $\mathscr{H}$ where $\mu^{A'} = 0$, and then displace away along $\lambda^A \pi^{A'}$. We want this displacement to work also for the *other* points on the α-curve. This is O.K. in flat space since we can take the direction of $\lambda^A \pi^{A'}$ elsewhere on the α-curve to be *constant*.

In curved space it seems possible that the λ^A that we have already been using might still work (i.e., $\underset{x}{\lambda}^A = \lambda^A$), but this $\pi^{A'}$ is *not* correct (i.e., $\underset{x}{\pi}^{A'} \neq \pi^{A'}$, in general). We seem to need a modification of $\pi^{A'}$, in order for this to work elsewhere along the curve. So far we do not have a suitable primed spinor that varies along the α-curve in a way that is not proportional to $\pi^{A'}$. We need all this so that we can propagate (W_α, Z^α) when $W_\alpha Z^\alpha \neq 0$, and there is no null geodesic defined by (W_α, Z^α).

More work is required.

References

LeBrun, C.R. (1986) Thickenings and gauge fields, Class. & Quant. Grav., **3**, 1039-59.

Penrose, R. and Rindler, W. (1986) *Spinors and Space-Time*, Vol. 2, Cambridge University Press, Cambridge.

Woodhouse, N.M.J. (1985) Real methods in twistor theory, Class. & Quant. Grav., **2**, 257-91.

Document generated by Anna’s Archive around 2023-2024 as part of the DuXiu collection (https://annas-blog.org/duxiu-exclusive.html).

Images have been losslessly embedded. Information about the original file can be found in PDF attachments. Some stats (more in the PDF attachments):

{
  "filename": "NDA3ODg1Mzguemlw",
  "filename_decoded": "40788538.zip",
  "filesize": 141801657,
  "md5": "ebebff60ee06b54832a10a9b85bcb287",
  "header_md5": "cc17074e6c4ada16e1816ca90755a4c2",
  "sha1": "fe549f2ce4fcd639802d1f4560211ba63a9f6c25",
  "sha256": "e1ab65606ebead12f1380104c2270f298b612636c343f2e8fb8b7f207f91497f",
  "crc32": 844337086,
  "zip_password": "",
  "uncompressed_size": 179725050,
  "pdg_dir_name": "ROGER PENROSE COLLECTED WORKS VOLUME 4 1981-1989_40788538",
  "pdg_main_pages_found": 859,
  "pdg_main_pages_max": 859,
  "total_pages": 884,
  "total_pixels": 5433450288,
  "pdf_generation_missing_pages": false
}